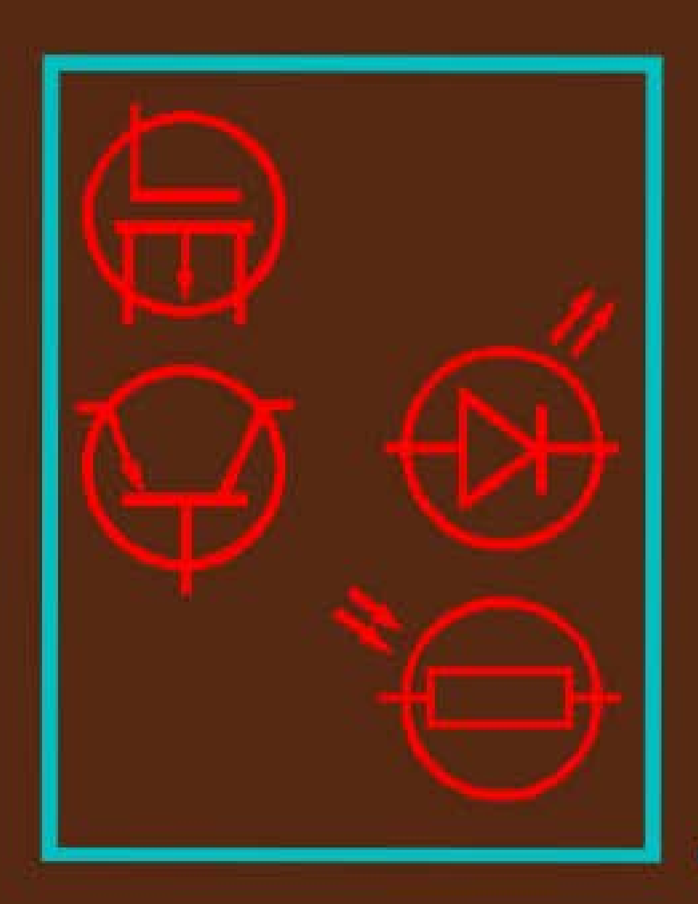
В. В. ПАСЫНКОВ
Л. К. ЧИРКИН

ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ



учебник для вузов

ББК 32.852
П 19
УДК 621.382

Оглавление

Рецензент — кафедра «Полупроводниковые приборы» Московского института радиотехники, электроники и автоматики (зав. кафедрой доц. Н. В. Короткова)

Пасынков В. В., Чиркин Л. К.

П 19 Полупроводниковые приборы: Учеб. для вузов по спец. «Полупроводники и диэлектрики» и «Полупроводниковые и микроэлектронные приборы» — 4-е изд., перераб. и доп. — М.: Высш. шк., 1987. — 479 с. ил.

В книге рассмотрены физические процессы в полупроводниковых приборах и элементах интегральных микросхем, их основные свойства, характеристики и параметры, конструктивно-технологические особенности полупроводниковых приборов в интегральном исполнении и общие принципы микроэлектроники.

В 4-м издании (3-е — 1981 г.) переработан материал, относящийся к диодам, тиристорам и другим приборам.

П $\frac{2403000000-196}{001(01)-87}$ 197—87

ББК 32.852
6ФО.032

ПРЕДИСЛОВИЕ

Инженер электронной техники по специальностям 0604 «Полупроводники и диэлектрики» и 0629 «Полупроводниковые и микроэлектронные приборы» должен быть специалистом по проектированию, конструированию, технологии и применению приборов и устройств, основанных на различных физических процессах в твердом теле. Без зиания принципа действия и свойств коикретного полупроводникового прибора невозможно правильно выполнить расчет, разработать технологию изготовления и организовать производство, исследовать свойства и измерить параметры этого прибора, а также рационально использовать этот прибор в той или иной схеме, в той или иной установке при различных условиях эксплуатации. Соответственно без прочного усвоения физики полупроводниковых приборов и элементов интегральных микросхем невозможно понять и усвоить даже основные положения практически всех специальных дисциплин, которые студенты будут изучать в соответствии с учебными планами после дисциплин «Полупроводниковые приборы и микроэлектроника», «Электронные приборы и микроэлектроника», «Физика полупроводниковых приборов», «Электронные и полупроводниковые приборы» и др.

Книга может быть использована не только в качестве учебника для отмечеиных специальностей, но и в качестве учебного пособия для студентов других специальностей электронной техники. Поэтому в гл. 1 книги изложены основиые сведения по физике полупроводников в минимальном объеме, необходимом для понимания принципа действия и свойств различных полупроводниковых приборов и элементов интегральных микросхем.

Перед изучением полупроводниковых приборов следует ознакомиться с обозначениями физических величин, принятыми в книге. Принцип действия большиинства полупроводииковых приборов основан на использовании разнообразных свойств выпрямляющих электрических переходов. Поэтому для глубокого понимания свойств и характеристик различных полупроводниковых приборов необходимо усвоить материал гл. 2, посвященной контактным явлениям, и гл. 3, где рассмотрены полупроводниковые диоды, структура которых проще структур большинства других приборов. При рассмотрении конкретных видов полупроводнико-

вых приборов следует обращать внимаиие иа связь их свойств и особениостей с ранее изучеиными общими закоиомериостями.

Киига составлеиа иа осиове опыта чтеиия лекций авторами и их коллегами на кафедре диэлектриков и полупроводииков Ленинградского ордеиа Лениик и ордена Октябрьской Революции электротехиического института им. В. И. Ульянова (Ленина). Большую помощь авторам оказали замечания рецензеитов первых трех изданий учебиика — лауреата Ленинской премии проф. Я. А. Федотова, проф. К. В. Шалимовой, чл.-кор. АН БССР В. А. Лабунова и их сотрудников. В составление учебника виесли вклад и преподаватели многих вузов СССР, которые, повышая свою квалификацию при кафедре диэлектриков и полупроводников ЛЭТИ, прииимали участие в обсуждении различных вопросов физики полупроводииковых приборов и методики их изложения.

Глубокую благодарность выражаем преподавателям кафедры полупроводииковых приборов Московского иистнтута радиотехиики, электроники и автоматики, возглавляемой доц. Н. В. Коротковой, за ценные замечания, высказаииые при рецеизировании рукописи четвертого издания учебиика.

Авторы

ВВЕДЕНИЕ

В Осиовиых направлениях экономического и социального развития СССР на 1986—1990 годы и на период до 2000 года поставлеиа задача обеспечить поступательный рост экономики, настойчиво повышать эффективность производства на основе его всесторонией иитенсификации и обеспечить дальнейшее ускореиие иаучиотехиического прогресса.

Научио-технический прогресс немыслим без электроиики. Интенсивное развитие электроники связаио с появлением иовых разнообразиых полупроводниковых приборов и интегральных микросхем, которые иаходят широкое применение в вычислительной технике, автоматике, радиотехнике и телевидении, в установках измерительиой техиики, медицине, биологни н т. д.

Полупроводниковые приборы в виде точечных диодов, или, как их раньше называли, кристаллических детекторов, примеияли еще в первых электронных установках. Выпрямительные свойства контактов между металлами и иекоторыми сериистыми соединеииями были обнаружены в 1874 г. В 1895 г. А. С. Поповым при изобретении радио был применен порошковый когерер, в котором использовались нелииейиые свойства зернистых систем. В 1922 г. О. В. Лосев использовал отрицательное дифференциальное сопротивление, возникающее при определениых условиях на точечиых контактах металла с полупроводником, для генерации и усиления высокочастотных электромагнитиых колебаний. Кроме того, им было обиаружеио свечение кристаллов карбида кремиия при прохождеиии тока через точечный контакт.

Однако в этот период успешно развивается техника электровакуумных приборов и из-за недостаточного зиаиия строения полупроводииков и происходящих в иих электрофизических процессов полупроводииковые приборы не получилн тогда существенного развития и примеиения.

В годы Великой Отечественной войны были разработаны точечные высокочастотные н сверхвысокочастотиые гермаииевые и кремииевые диоды. В 1942 г. в СССР был начат выпуск полупроводниковых термоэлектрических геиераторов для непосредственного преобразования тепловой энергии в электрическую. Термогенераторы использовались для питания переносных радиостанций партизанских отрядов. Создание и производство этих и

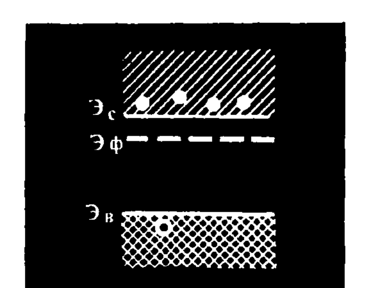
многих других приборов стало возможным благодаря фундаментальным теоретическим и экспериментальным исследованиям свойств полупроводников, проведенным группой ученых под руководством академика А. Ф. Иоффе.

С 1948 г., т. е. со времени создания американскими учеными Дж. Бардином, В. Браттейном и В. Шокли точечного транзистора, начался новый этап развития полупроводниковой электроники. В 50-х годах были разработаны различные типы биполярных транзисторов, тиристоров, мощных германиевых и кремниевых выпрямительных диодов, фотодиодов, фототранзисторов, кремниевых фотоэлементов, туннельных диодов и др.

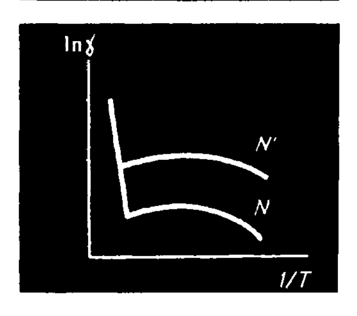
Принцип работы полевых транзисторов с изолированным затвором был предложен еще в 30-х годах нашего столетия, но до окончательной разработки этих транзисторов потребовались многолетние исследования электрофизических процессов на границе полупроводника с диэлектриком и технологии необходимых структур.

В 60-х годах нашего столетия начинается производство интегральных микросхем. При этом удалось существенно уменьшить стоимость и повысить надежность устройств полупроводниковой электроники, значительно уменьшить их массу и габаритные размеры формированием всех пассивных и активных элементов интегральных микросхем в едином технологическом процессе, а также в результате конструктивной интеграции.

Современный этап развития полупроводниковой электроники характеризуется в нашей стране большим объемом научно-исследовательскнх и технологических работ, направленных на дальнейшее совершенствование имеющихся и создание новых полупроводниковых приборов и интегральных микросхем.



1



Основные сведения по физике полупроводников

§ 1.1. ЭНЕРГЕТИЧЕСКИЕ ЗОНЫ ПОЛУПРОВОДНИКОВ

Полупроводник — вещество, основным свойством которого является сильная зависимость удельной проводимости от воздействия внешних факторов (температуры, электрического поля, света и др.).

Каждый электрон, входящий в состав атома, обладает определенной полной энергией или занимает определенный энергетический уровень.

В твердом теле благодаря взаимодействию атомов энергетические уровни расщепляются и образуют *энергетические зоны*, состоящие из отдельных близко расположенных по энергии уровней, число которых соответствует числу однородных атомов в данном кристаллическом теле (рис. 1.1). Энергетическую зону или совокупность нескольких перекрывающихся энергетических зон, которые образовались в результате расщепления одного или нескольких энергетических уровней отдельного атома, называют *разрешенной зоной*. Электроны в твердом теле могут иметь энергии, соответствующие разрешенной зоне. Верхний энергетический уровень разрешенной зоны называют *потолком*, нижний — *дном*.

Энергетические уровни валентных электронов при расщеплении образуют *валентную зону*. Разрешенные энергетические уровни, свободные от электронов в невозбужденном состоянии атома, расщепляясь, образуют одну или несколько *свободных зон*. Нижнюю из свободных зон называют *зоной проводимости*.

Наибольший интерес представляют валентная зона и зона проводимости, так как от их взаимного расположения и от степени их заполнения электронами зависят электрические, оптические и другие свойства твердых тел. Между разрешенными зонами находятся *запрещенные зоны*, т. е. области значений энергии,

которыми не могут обладать электроны в идеальном кристалле. Для полупроводников (согласно сказанному) наибольшее значение имеет запрещенная зона, разделяющая валентную зону и зону проводимости. Она характеризуется *шириной запрещенной зоны* $\Delta Э$, т. е. разностью энергий дна зоны проводимости и потолка валентной зоны.

При температуре 300 К у кремния ширина запрещенной зоны $\Delta Э = 1{,}12$ эВ; у германия $\Delta Э = 0{,}75$ эВ; у арсенида галлия $\Delta Э = 1{,}43$ эВ; у карбида кремния $\Delta Э = 2{,}4 \div 3{,}4$ эВ (для разных политипов).

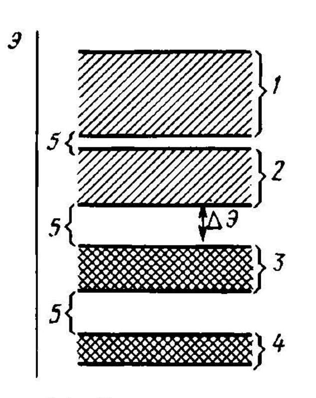


Рис. 1.1. Энергетические зоны полупроводника:

1, 2, 3, 4 — разрешенные зоны; *5* — запрещенные зоны; *1, 2* — свободные зоны; *2* — зона проводимости; *3* — валентная зона; $\Delta Э$ — ширина запрещенной зоны

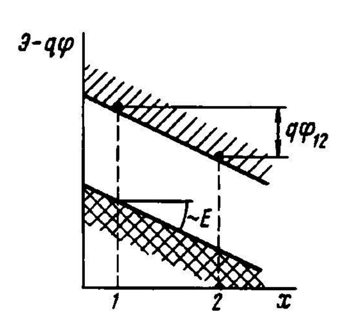


Рис. 1.2. Энергетическая диаграмма полупроводника при наличии в нем электрического поля напряженностью E и разности потенциалов между точками *1* и *2*, равной φ_{12}

Ширина запрещенной зоны изменяется с изменением температуры. Происходит это в результате: 1) изменения амплитуды тепловых колебаний атомов кристаллической решетки; 2) изменений межатомных расстояний, т. е. объема тела. С повышением температуры в первом случае ширина запрещенной зоны уменьшается, во втором случае может быть как уменьшение, так и увеличение ширины запрещенной зоны. У большинства полупроводников ширина запрещенной зоны с повышением температуры уменьшается.

При наличии в полупроводнике электрического поля энергетические диаграммы целесообразно строить, откладывая по вертикальной оси полную энергию электронов $Э - q\varphi$ (с учетом потенциальной энергии электрона в электрическом поле), а по горизонтальной оси — геометрическую координату (рис. 1.2). При таком построении энергетических диаграмм в областях, где существует электрическое поле, энергетические уровни и зоны получаются наклонными, причем угловой коэффициент пропорционален на-

пряженности электрического поля с учетом принятых масштабов по осям, а относительное смещение соответствующих энергетических уровней или зон — разности потенциалов между данными точками объема полупроводника. Коэффициент пропорциональности при этом равен элементарному заряду электрона q; увеличению потенциала соответствует понижение энергетических уровней или зон на энергетической диаграмме.

§ 1.2. ГЕНЕРАЦИЯ И РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА

Образование свободных электронов и дырок — *генерация носителей заряда* — происходит при воздействии теплового хаотического движения атомов кристаллической решетки (*тепловая генерация*), при воздействии поглощенных полупроводником квантов света (*световая генерация*) и других энергетических факторов. Так как полупроводник всегда находится под действием всех этих факторов или хотя бы одного ($T \neq 0$), генерация носителей происходит непрерывно.

Одновременно с генерацией в полупроводнике непрерывно происходит и обратный процесс — *рекомбинация носителей заряда,* т. е. возвращение электронов из зоны проводимости в валентную зону, в результате чего исчезает пара носителей заряда. В состоянии термодинамического равновесия процессы генерации и рекомбинации носителей заряда взаимно уравновешены. При этом в полупроводнике существуют *равновесные концентрации электронов* n_0 и *дырок* p_0.

При воздействии на полупроводник нетеплового внешнего энергетического фактора (света, сильного электрического поля и др.) из-за генерации новых носителей заряда их концентрация n и p (*неравновесная концентрация*) будет превышать равновесную концентрацию на величину Δn (или Δp), которую называют *избыточной концентрацией.* Таким образом,

$$\Delta n = n - n_0; \quad \Delta p = p - p_0. \tag{1.1}$$

Избыточная концентрация носителей заряда может возникать в отдельных областях полупроводниковой структуры прибора не только в результате внешних энергетических воздействий, но и за счет различных процессов (инжекции, экстракции, аккумуляции и т. д.), которые могут происходить в полупроводниковых приборах (о чем говорится далее).

Механизмы рекомбинации могут быть различны (рис. 1.3).

Межзонная, или непосредственная, рекомбинация происходит при переходе свободного электрона из зоны проводимости в валентную зону на один из свободных энергетических уровней, что соответствует исчезновению пары носителей заряда — свободного электрона и дырки. Однако такой процесс межзонной рекомбинации маловероятен, так как свободный электрон и дырка должны оказаться одновременно в одном и том же месте крис-

талла. Кроме того, должен выполняться закон сохранения импульса, т. е. рекомбинация электрона и дырки возможна только при одинаковых, но противоположно направленных импульсах электроиа и дырки. Поэтому, например, в германии на 10 тыс. рекомбинаций лишь одна происходит в результате межзонной рекомбииации.

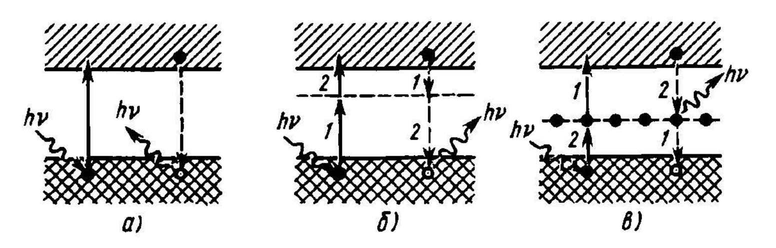


Рис. 1.3. Различные механизмы генерации и рекомбинации носителей заряда:

а — межзонная генерация и рекомбинация; *б* — генерация и рекомбинация с участием пустых рекомбинационных ловушек; *в* — генерация и рекомбинация с участием заполненных электронами рекомбинационных ловушек; ———— — условное обозначение генерации; — — — — условное обозначение рекомбинации. Цифры означают этапы процессов генерации и рекомбинации

Рекомбинация с участием рекомбинационных ловушек протекает в два этапа. На первом этапе рекомбинационная ловушка (или энергетический уровень рекомбинационной ловушки) захватывает, например, электрон из зоны проводимости. Таким образом, электрон выбывает из процесса электропроводности. В этом состоянии ловушка будет находиться до тех пор, пока к ней не подойдет дырка, или, другими словами, пока в данном месте кристалла не окажется свободный энергетический уровень валентной зоны. При выполнении этих условий осуществляется второй этап рекомбинации — электрон переходит на свободный уровень валентной зоны (что эквивалентно захвату дырки из валентной зоны отрицательно заряженной ловушкой).

Двухэтапный процесс рекомбинации более вероятен, так как он не требует одновременного присутствия в данном месте кристалла свободного электрона и дырки. Рекомбинационная ловушка воспринимает количество движения, необходимое для соблюдения закона сохранения импульса, и может забрать часть энергии, освобождаемой в процессе рекомбинации.

Роль рекомбинационных ловушек могут выполнять примесные атомы или ионы, различные включения в кристалле, незаполненные узлы кристаллической решетки, трещины и другие несовершенства объема или поверхности.

В связи с тем что на поверхности кристалла перечисленных дефектов значительно больше, чем в объеме, процесс рекомбинации на поверхности должен идти значительно интенсивнее. Его рассматривают и оценивают обычно отдельно, считая *поверх-*

ностную рекомбинацию разновидностью рекомбинации с участием рекомбиинационных ловушек.

В зависимости от того, как расходуется энергия, освобождающаяся при рекомбинации электрона и дырки, рекомбинацию можно подразделить на два вида.

Излучательной рекомбинацией называют рекомбинацию, при которой энергия, освобождающаяся при переходе электрона на более низкий энергетический уровень, излучается в виде кванта света (фотона).

При *безызлучательной (фононной) рекомбинации* избыточная энергия электрона передается кристаллической решетке полупроводника, т. е. избыточная энергия идет на образование фононов — квантов тепловой энергии.

§ 1.3. КОНЦЕНТРАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКЕ ПРИ ТЕРМОДИНАМИЧЕСКОМ РАВНОВЕСИИ

В соответствии со статистикой Фéрми—Дирáка вероятность заполнения энергетического уровня электроном определяется энергией $Э$, соответствующей этому уровню, и абсолютной температурой T:

$$P_n(Э) = \frac{1}{1 + \exp\frac{Э - Э_ф}{kT}}, \qquad (1.2)$$

где $Э_ф$ — энергия уровня Ферми, вероятность заполнения которого равна 1/2 и относительно которого кривая вероятности симметрична (рис. 1.4).

Энергия *уровня Ферми* соответствует верхней границе электронного распределения при температуре $T = 0$, а также средней энергии «диапазона размытия» при любой другой температуре (рис. 1.4). Симметрия кривой вероятности заполнения относительно уровня Ферми означает одинаковую вероятность заполнения уровня электроном с энергией, большей на величину $Э - Э_ф$, и вероятность освобождения уровня от электрона с энергией, на столько же меньшей энергии уровня Ферми.

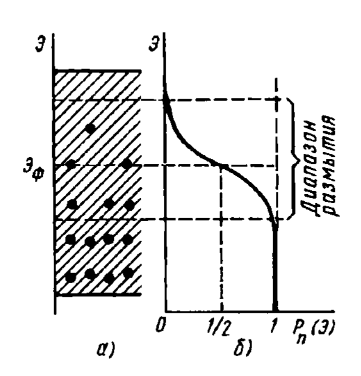


Рис. 1.4. Распределение электронов в частично заполненной зоне (*а*) и фукнция вероятности заполиения энергетических уровней (*б*)

С помощью соотношения (1.2) можно определять заполнение электронами зоны проводимости или валентной зоны полупроводника. Но для валентной зоны удобнее гово-

рить о дырках — пустых энергетических уровнях в валентиой зоне. Любой энергетический уровень может либо быть заият электроном, либо свободен от электрона. Поэтому сумма вероятностей этих двух событий должна быть равна единице: $P_n(Э) + P_p(Э) = 1$.

Тогда вероятность заполнения энергетического уровня дыркой

$$P_p(Э) = \frac{1}{1 + \exp\left(-\frac{Э - Э_ф}{kT}\right)}. \qquad (1.3)$$

Уровень Ферми обычно расположен в запрещенной зоне энергетической диаграммы относительно далеко (в единицах энергии) от зоны проводимости и от валентной зоны по сравнению с энергией kT (при комнатной температуре $kT \approx 0{,}025$ эВ), т. е.

$$|Э - Э_ф| \gg kT. \qquad (1.4)$$

Поэтому, пренебрегая еднницей в знаменателе (1.2), вероятность распределения электронов по энергетическим уровням зоны проводимости определим с помощью статистики Максвелла — Больцмана:

$$P_n(Э) \approx \exp\left(-\frac{Э - Э_ф}{kT}\right) \qquad (1.5)$$

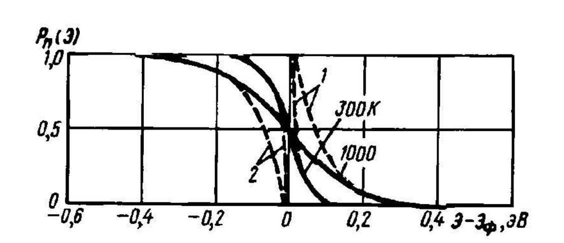


Рис. 1.5. Вероятность заполнения электронами энергетических уровней при различных температурах:

——— — по статистнке Ферми—Дирака; — — — — по статистике Максвелла—Больцмана для электронов в зоне проводимости *1* и в валеитной зоне *2*

Аналогично найдем вероятность распределения дырок по энергетическим уровням валентной зоны с учетом (1.3) и (1.4):

$$P_p(Э) \approx \exp\frac{Э - Э_ф}{kT}. \qquad (1.6)$$

Таким образом, для большинства полупроводников (*невырожденных*) можно пользоваться статистикой Максвелла — Больцмана и только в некоторых случаях для полупроводников (*вырожденных*) необходимо использовать статистику Ферми—Дирака. Разница в этих двух функциях распределения электронов по энергиям показана на рис. 1.5.

Для определения концентрации электронов в невырожденном полупроводнике надо проинтегрировать по энергии произведение удвоенной функции распределения плотности энергетических уровней в зоне проводимости $[2N(Э)]$ и вероятности заполнения этих уровней электронами (1.5). Интегрирование нужно проводить от энергии дна до энергии потолка зоны проводимости. Если же учесть ничтожно малую вероятность заполнения уровней электронами у потолка зоны проводимости, то можно верх-

ний предел интегрирования считать равным бесконечности, т. е.

$$n_0 = \int_{Э_c}^{\infty} 2N(Э) \cdot P_n(Э) dЭ.$$

В результате интегрирования получим

$$n_0 = 2N_c \exp\left(-\frac{Э_c - Э_ф}{kT}\right), \quad (1.7)$$

где $2N_c$ — эффективная плотность состояний в зоне проводимости; N_c — эффективная плотность энергетических уровней в зоне проводимости, энергия которых приведена ко дну зоны проводимости.

Эффективная плотность состояний отличается в два раза от эффективной плотности уровней в связи с тем, что на каждом энергетическом уровне могут находиться два электрона с различными спинами (по принципу Паули).

Аналогично, равновесная концентрация дырок в любом невырожденном полупроводнике при термодинамическом равновесии

$$p_0 = 2N_в \exp \frac{Э_в - Э_ф}{kT}, \quad (1.8)$$

где $N_в$ — эффективная плотность энергетических уровней в валентной зоне, энергия которых приведена к потолку валентной зоны; $Э_в$ — энергия потолка валентной зоны.

Таким образом, равновесная концентрация носителей заряда определяется эффективной плотностью уровней в соответствующей зоне, температурой и положением уровня Ферми относительно дна зоны проводимости или потолка валентной зоны.

§ 1.4. СОБСТВЕННЫЕ ПОЛУПРОВОДНИКИ

Собственный полупроводник — это полупроводник без донорных и акцепторных примесей или с концентрацией примеси настолько малой, что она не оказывает существенного влияния на удельную проводимость полупроводника. В собственном полупроводнике при температуре абсолютного нуля отсутствуют носители заряда, так как валентная зона полностью занята электронами (там нет дырок), а в зоне проводимости нет электронов.

При температурах выше абсолютного нуля некоторые электроны валентной зоны могут быть переброшены в зону проводимости — возможна тепловая генерация пар носителей заряда, в зоне проводимости появляются свободные электроны, а в валентной зоне — дырки. Процесс тепловой генерации возможен даже при очень низких температурах из-за значительных флуктуаций (отклонений) энергий тепловых колебаний атомов от средней энергии тепловых колебаний атомов относительно узлов кристаллической решетки.

Как отмечалось, кроме тепловой генерации носителей заряда в полупроводнике существует и их рекомбинация, и эти процессы при любой температуре взаимно уравновешены. При этом в собственном полупроводнике существует *собственная концентрация электронов,* которая может быть выражена соотношением, аналогичным (1.7):

$$n_i = \sqrt{4N_c N_в} \exp\left(-\frac{Э_c - Э_Ф}{kT}\right) \approx 2N_c \exp\left(-\frac{Э_c - Э_Ф}{kT}\right). \quad (1.9)$$

Формула для *собственной концентрации дырок* аналогична (1.8):

$$p_i = \sqrt{4N_c N_в} \exp\frac{Э_в - Э_Ф}{kT} \approx 2N_в \exp\frac{Э_в - Э_Ф}{kT}. \quad (1.10)$$

Эффективные плотности уровней в зоне проводимости и в валентной зоне определяются соотношениями

$$N_c = \left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2}, \qquad N_в = \left(\frac{2\pi m_p^* kT}{h^2}\right)^{3/2}, \quad (1.11)$$

где m_n^* и m_p^* — эффективные массы электронов и дырок (смысл этих параметров рассмотрен далее); h — постоянная Планка.

Так как в собственном полупроводнике носители заряда образуются в результате ионизации собственных атомов полупроводника, т. е. благодаря перебросу электронов из валентной зоны в зону проводимости, то одновременно возникает два носителя заряда противоположных знаков. Поэтому

$$n_i = p_i. \quad (1.12)$$

Определим положение уровня Ферми для собственного полупроводника. Учитывая соотношения (1.9) — (1.12), получим

$$(m_n^*)^{3/2} \exp\left(-\frac{Э_c - Э_Ф}{kT}\right) = (m_p^*)^{3/2} \exp\frac{Э_в - Э_Ф}{kT},$$

откуда

$$Э_Ф = \frac{Э_c + Э_в}{2} + \frac{3}{4} kT \ln\frac{m_p^*}{m_n^*} \quad (1.13)$$

При небольшом различии эффективных масс электронов и дырок

$$Э_Ф \approx \frac{Э_c + Э_в}{2} = Э_i, \quad (1.14)$$

т. е. в собственном полупроводнике приближенно можно считать, что уровень Ферми расположен посередине запрещенной зоны.

Приняв в данном случае для простоты начало отсчета энергии от потолка валентной зоны, т. е. $Э_в = 0$, собственные кон-

центрации носителей можно выразить через ширину запрещенной зоны $\Delta Э = Э_с - Э_в$:

$$n_i = 2N_c \exp\left(-\frac{\Delta Э}{2kT}\right); \qquad (1.15)$$

$$p_i = 2N_в \exp\left(-\frac{\Delta Э}{2kT}\right). \qquad (1.16)$$

Из соотношений (1.15) и (1.16) можно определить собственные концентрации носителей в разных полупроводниках. При комнатной температуре ($T \approx 300$ К) в кремнии $n_i \approx 10^{10}$ см$^{-3}$, в германии $n_i \approx 10^{13}$ см$^{-3}$.

§ 1.5. ПРИМЕСНЫЕ ПОЛУПРОВОДНИКИ

Большинство полупроводниковых приборов изготовляют на основе *примесных полупроводников,* т. е. полупроводников, электрические характеристики которых определяются донорными или акцепторными примесями. Таким образом, в рабочем диапазоне температур полупроводникового прибора поставщиками основного количества носителей заряда в полупроводниковом материале должны быть примеси. Поэтому в практике важное значение имеют такие полупроводниковые материалы, у которых ощутимая собственная концентрация носителей заряда появляется при возможно более высокой температуре, т. е. полупроводники с достаточно большой шириной запрещенной зоны.

В элементарных полупроводниках, состоящих в основном из атомов одного химического элемента, примесями являются чужеродные атомы. В полупроводниковых соединениях и твердых растворах, состоящих из атомов двух или большего числа химических элементов, примесями могут быть не только включения атомов посторонних химических элементов, но и избыточные по отношению к стехиометрическому составу атомы химических элементов, входящих в химическую формулу сложного полупроводника.

Такую же роль, как примеси, могут играть различные дефекты кристаллической решетки: пустые узлы, дислокации или сдвиги, возникающие при пластической деформации кристалла, и т. д.

Рассмотрим роль примесей, атомы которых создают дискретные энергетические уровни в пределах запрещенной зоны полупроводника. При небольшой концентрации примесей их атомы расположены в полупроводнике на таких больших расстояниях друг от друга, что не взаимодействуют между собой. Поэтому нет расщепления примесных уровней. Вероятность непосредственного перехода электронов от одного примесного атома к другому ничтожно мала, т. е. с точки зрения зонной теории ничтожно мала вероятность перехода электрона с одного дискретного примесного уровня на другой.

При большой концентрации примесей в результате взаимодействия примесных атомов между собой примесные уровни одного типа расщепляются в энергетическую *примесную зону*. Электроны, находящиеся в примесной зоне, так же как в зоне проводимости и в валентной зоне при неполном их заполнении, могут переходить с уровня на уровень, приобретая необходимую для этого энергию за счет ускорения во внешнем электрическом поле на длине свободного пробега.

Примеси могут быть донорного и акцепторного типа.

Донор — это примесный атом или дефект кристаллической решетки, создающий в запрещенной зоне энергетический уровень, занятый в невозбужденном состоянии электроном и способный в возбужденном состоянии отдать электрон в зону проводимости.

Акцептор — это примесный атом или дефект кристаллической решетки, создающий в запрещенной зоне энергетический уровень, свободный от электрона в невозбужденном состоянии и способный захватить электрон из валентной зоны в возбужденном состоянии.

Доноры и акцепторы в полупроводнике могут быть ионизированы под действием энергии, поступающей в кристалл в виде квантов света, теплоты и т. д.

Под *энергией ионизации донора* понимают минимальную энергию, которую необходимо сообщить электрону, находящемуся на донорном уровне, чтобы перевести его в зону проводимости.

Энергия ионизации акцептора — это минимальная энергия, которую необходимо сообщить электрону валентной зоны, чтобы перевести его на акцепторный уровень.

Энергия ионизации примесных атомов значительно меньше энергии ионизации собственных атомов полупроводника или ширины запрещенной зоны. Поэтому в примесных полупроводниках при низких температурах преобладают носители заряда, возникшие из-за ионизации примесей. Если электропроводность полупроводника обусловлена электронами, его называют *полупроводником n-типа*, если электропроводность обусловлена дырками — *полупроводником p-типа*.

Обычно в полупроводниках присутствуют как доноры, так и акцепторы. Полупроводник, у которого концентрация доноров равна концентрации акцепторов, называют *скомпенсированным*.

Соотношения для концентрации электронов (1.7) и дырок (1.8) можно преобразовать, приняв $Э_в = 0$ и учтя соотношения (1.15) и (1.16):

$$n_0 = 2N_c \exp\left(-\frac{Э_i + Э_i - Э_Ф}{kT}\right) = n_i \exp\left(-\frac{Э_i - Э_Ф}{kT}\right); \quad (1.17)$$

$$p_0 = 2N_в \exp\frac{Э_в - Э_Ф - Э_i + Э_i}{kT} = p_i \exp\frac{Э_i - Э_Ф}{kT}. \quad (1.18)$$

Отсюда следует, что

$$n_0 p_0 = n_i^2, \quad (1.19)$$

т. е. в невырожденном полупроводнике произведение концентраций свободных электронов и дырок при термодинамическом равновесии есть постоянная величина, равная квадрату собственной концентрации при данной температуре.

Соотношение (1.19) справедливо для невырожденного полупроводника, т. е. для полупроводника, у которого уровень Ферми расположен в запрещенной зоне достаточно далеко (на 2—3 kT) от дна зоны проводимости или от потолка валентной зоны, так как только при этих условиях можно пользоваться функцией распределения Максвелла — Больцмана.

Соотношение (1.19) обычно называют *законом действующих масс* в соответствии с терминологией химической термодинамики (константа химического равновесия выводится из закона действующих масс).

§ 1.6. ВРЕМЯ ЖИЗНИ НЕРАВНОВЕСНЫХ НОСИТЕЛЕЙ ЗАРЯДА

После прекращения энергетического воздействия на полупроводник избыточная концентрация носителей заряда в нем из-за процесса рекомбинации через некоторое время уменьшится до нуля. Количество носителей заряда, рекомбинирующих в единицу времени в единице объема (быстрота изменения концентрации), пропорционально избыточной концентрации и обратно пропорционально некоторому параметру τ, который называют временем жизни:

$$-\frac{dn}{dt} = -\frac{d(\Delta n)}{dt} = \frac{\Delta n}{\tau_n}, \quad -\frac{dp}{dt} = -\frac{d(\Delta p)}{dt} = \frac{\Delta p}{\tau_p}. \qquad (1.20)$$

Таким образом, *временем жизни неравновесных носителей заряда* является отношение избыточной концентрации (Δn или Δp) неравновесных носителей заряда к скорости изменения этой концентрации вследствие рекомбинации:

$$\tau_n = \frac{|\Delta n|}{|d(\Delta n)/dt|} \text{ и } \tau_p = \frac{|\Delta p|}{|d(\Delta p)/dt|}. \qquad (1.21)$$

Рассмотрим наиболее простой частный случай — постоянство концентрации носителей заряда, с которыми происходит рекомбинация неравновесных носителей. Этот случай реализуется, например, в полупроводнике с явно выраженной примесной электропроводностью при введении в него неосновных носителей заряда в небольшом количестве. Тогда появление неравновесных неосновных носителей заряда не вызывает существенного изменения концентрации основных, с которыми происходит рекомбинация. Время жизни при этом оказывается постоянным, а количество носителей заряда, рекомбинировавших в единицу времени в единице объема, пропорционально первой степени избыточной кон-

центрации. Этот случай называют *линейной рекомбинацией.* Уравнения (1.20) решаются очень просто.

Пусть в момент времени $t = 0$ в полупроводнике создаиа избыточная концентрация $\Delta n(0)$ (или $\Delta p\,(0)$), которая после окончания действия источника избыточных носителей заряда должиа стремиться к нулю при $t \to \infty$. Решение уравнений (1.20) при этих условиях имеет вид

$$\Delta n(t) = \Delta n(0) \exp(-t/\tau_n) \text{ и } \Delta p(t) = \Delta p(0) \exp(-t/\tau_p),$$

т. е. при лииейной рекомбинации избыточная концентрация носителей за время жизни уменьшается в e раз.

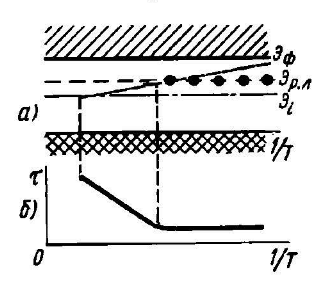


Рис. 1.6. Пояснение температурной зависимости времени жизни:

а — завнснмость положения уровня Ферми от температуры и заполнение уровней рекомбинационных ловушек электронамн; *б* — завнснмость времени жнзнн от температуры

Время жизни неравновесных носителей заряда зависит от температуры полупроводника. Рассмотрим температурную зависимость времени жизни на примере полупроводника с электропроводностью n-типа с рекомбинационными ловушками в верхней половине запрещенной зоны (рис. 1.6). При очень низкой температуре рекомбинационные ловушки заполнены электронами, так как вероятность их заполнения, судя по положению уровня Ферми, намного больше 50%. При этом первый этап рекомбинации (захват дырки рекомбинационной ловушкой) происходит быстро и время жизни оказывается небольшим. С повышением температуры уровень Ферми смещается вниз и находится вблизи энергетических уровней рекомбинационных ловушек. Это означает, что теперь не все ловушки заполнены электронами, т. е. не все ловушки могут захватить блуждающие по полупроводнику дырки. Поэтому с повышением температуры время жизни растет.

Необходимо отметить, что рассмотренная температурная зависимость времени жизни справедлива только для полупроводника с рекомбинационными ловушками одного сорта или типа. Если же в полупроводнике будут рекомбинационные ловушки разных типов, создающие в запрещенной зоне несколько различных энергетических уровней, то температурная зависимость времени жизни может быть сложнее представленной на рис. 1.6.

Время жизни носителей заряда в значительной степени зависит от концентрации в полупроводнике рекомбинационных ловушек, так как рекомбинация чаще всего происходит с их помощью. Но кроме рекомбинационных ловушек в запрещенной зоне полупроводника обычно существуют энергетические уровни, которые могут захватывать носители только какого-либо одного типа. Такие уровни называют *уровнями ловушек захвата,* а де-

фекты кристаллической решетки, создающие уровни ловушек захвата, — *ловушками захвата.*

Энергетические уровни ловушек захвата электронов расположены в запрещенной зоне вблизи зоны проводимости, ловушек захвата дырок — вблизи валентной зоны (рис. 1.7).

Через некоторое время после захвата носителей заряда ловушкой захвата может произойти ионизация этой ловушки, т. е. освобождение носителя заряда. Если время нахождения носителя в ловушке захвата велико или велика концентрация ловушек захвата, то эффективное время жизни носителя заряда может оказаться значительно больше действительного времени жизни, так как находящийся в ловушке захвата носитель не может в это время рекомбинировать. Рекомбинация возможна только после ионизации ловушки захвата или после освобождения носителя заряда.

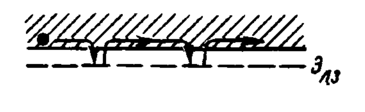


Рис. 1.7. Расположение энергетических уровней ловушек захвата и схематическое изображение процесса захвата электронов и дырок

Таким образом, в запрещенной зоне энергетической диаграммы полупроводника может существовать много различных локальных энергетических уровней, связанных с наличием разных примесей. Часть из них может быть уровнями ловушек захвата, часть — уровнями рекомбинационных ловушек. При различной степени отклонения от термодинамического равновесия роль, выполняемая отдельными ловушками, может изменяться, т. е. ловушки захвата могут стать рекомбинационными ловушками и наоборот. Чтобы установить количественный критерий отличия этих энергетических уровней, введены понятия *демаркационных уровней,* для которых вероятность ионизации с образованием носителя одного знака равна вероятности захвата носителя заряда противоположного знака.

В неравновесном состоянии распределение свободных электронов и дырок уже не соответствует распределению этих носителей заряда при термодинамическом равновесии. Поэтому в неравновесном состоянии распределение свободных электронов по энергетическим уровням характеризуется своим *квазиуровнем Ферми для электронов,* распределение дырок — своим *квазиуровнем Ферми для дырок,* которые имеют тот же смысл для полупроводника в неравновесном состоянии, что и уровень Ферми в условиях термодинамического равновесия. Чем больше неравновесные концентрации свободных электронов и дырок отклоняются от своих равновесных значений, тем больше отличается положение квазиуровней Ферми для электронов и для дырок от положения уровня Ферми в условиях термодинамического равновесия.

При качественном рассмотрении различных зависимостей параметров полупроводниковых приборов можно в первом прибли-

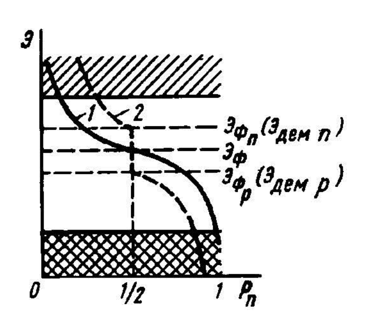


Рис. 1.8. Вероятность заполнения энергетических уровней в состоянии термодинамического равновесия (кривая *1*) и в неравновесном состоянии (кривая *2*) и расположение квазиуровней Ферми или демаркационных уровней для электронов и дырок

жении считать демаркационные уровни совпадающими с соответствующими квазиуровнями Ферми (рис. 1.8). Уровни, лежащие выше электронного демаркационного уровня или выше квазиуровня Ферми для электронов, являются уровнями захвата электронов. Вероятность их заполнения электронами меньше 50%, что соответствует большой вероятности (более 50%) переброса электронов с этих уровней в зону проводимости в результате тепловой генерации. Аналогично, уровни, лежащие ниже дырочного демаркационного уровня или квазиуровня Ферми для дырок, являются уровнями захвата дырок.

Для уровней, расположенных между электронными и дырочными демаркационными уровнями или между квазиуровнями Ферми, характерна большая вероятность заполнения, с одной стороны, электронами, а с другой стороны, дырками. В действительности сумма вероятностей заполнения какого-либо уровня электроном и дыркой должна быть равна единице. Поэтому следует считать, что заполнение электронами и дырками всех уровней, расположенных между демаркационными уровнями $Э_{дем\ n}$ и $Э_{дем\ p}$, одинаково. В связи с этим уровни, расположенные между электронным и дырочным демаркационными уровнями, следует считать уровнями рекомбинационных ловушек.

§ 1.7. ПРОЦЕССЫ ПЕРЕНОСА ЗАРЯДОВ В ПОЛУПРОВОДНИКАХ

Процесс переноса зарядов может наблюдаться в полупроводниках при наличии электронов в зоне проводимости и при неполном заполнении электронами валентной зоны. При выполнении этих условий и в отсутствие градиента температуры перенос носителей зарядов может происходить либо под действием электрического поля, либо под действием градиента концентрации носителей заряда.

Дрейф носителей заряда

Направленное движение носителей заряда под действием электрического поля называют *дрейфом*. Электроны, получая ускорение в электрическом поле, приобретают на длине свободного пробега дополнительную энергию около $10^{-8} - 10^{-4}$ эВ. При этом электроны переходят на более высокие энергетические уровни

(разница в энергиях между соседними энергетическими уровнями в разрешенной зоне около 10^{-22} эВ). При очередном соударении электрона с атомом кристаллической решетки электрон отдает кристаллической решетке накопленную на длине свободного пробега энергию, возвращаясь на один из низко лежащих энергетических уровней в разрешенной зоне. Так можно представить процесс электропроводности в результате движения электронов в зоне проводимости и в валентной зоне. Однако, учитывая почти полное заполнение электронами энергетических уровней валентной зоны, удобнее рассматривать в валентной зоне движение дырок: дырки, двигаясь по направлению вектора электрического поля и приобретая в этом поле дополнительную энергию, переходят по энергетической диаграмме на более низкие энергетические уровни.

В результате дрейфа электронов в полупроводнике появляется электронная составляющая плотности дрейфового тока, которую запишем на основании закона Ома:

$$\mathbf{J}_{n\,\text{др}} = \gamma_n \mathbf{E} = qn\mu_n \mathbf{E}, \qquad (1.22)$$

где γ_n — *удельная проводимость полупроводника* (любого материала) при одном виде имеющихся в нем носителей заряда — электронов; μ_n — *подвижность электронов,* т. е. величина, численно равная средней скорости их направленного движения в электрическом поле с напряженностью, равной единнце.

Аналогично, дырочная составляющая плотности дрейфового тока

$$\mathbf{J}_{p\,\text{др}} = \gamma_p \mathbf{E} = qp\mu_p \mathbf{E}. \qquad (1.23)$$

Знаки в правых частях соотношений (1.22) и (1.23) одинаковые, так как заряды электронов и дырок различны по знаку и движутся в электрическом поле в разные стороны.

Полная плотность дрейфового тока при наличии свободных электронов и дырок равна сумме электронной и дырочной составляющих:

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p = q\gamma \mathbf{E} = q(n\mu_n + p\mu_p)\mathbf{E}, \qquad (1.24)$$

где γ — удельная проводимость полупроводника (любого материала) с учетом свободных электронов и дырок.

Диффузия носителей заряда

Поведение свободных электронов и дырок в полупроводнике напоминает поведение молекул газа. Эту аналогию можно распространить и на явления, происходящие в результате неравномерного распределения концентрации носителей заряда в объеме полупроводника в отсутствие градиента температуры. В этом случае происходит *диффузия* — движение носителей заряда из-за градиента концентрации, выравнивание концентрации носителей по полупроводнику.

Из молекулярной физики известно, что поток частиц при диффузии (число частиц, пересекающих в единицу времени единичиую площадку, перпендикулярную направлению градиента концентрации) пропорционален градиенту концентрации этих частиц:

$$\Phi_m = -D_m \operatorname{grad} m, \qquad (1.25)$$

где D_m — *коэффициент диффузии*, равный абсолютному значению отношения плотности потока частиц к градиенту их концентрации.

Вектор граднента концентрации направлен в сторону возрастания аргумента, а частицы диффундируют туда, где их меньше, т. е. против градиента концентрации. Поэтому знаки перед правой и левой частями соотношения (1.25) различны.

Остановимся на процессе диффузии электронов и дырок в полупроводниках, т. е. на диффузии заряженных частиц (или квазичастиц). Так как всякое направленное движение одноименно заряженных частиц есть электрический ток, то плотность электронной составляющей диффузионного тока может быть получена путем умножения правой части (1.25) на элементарный заряд электрона. Электроны диффундируют против вектора градиента концентрации и имеют отрицательный заряд. Поэтому направление вектора плотности диффузионного тока электронов должно совпадать с направлением вектора градиента концентрации электронов, т. е.

$$\mathbf{J}_{n\,\text{диф}} = qD_n \operatorname{grad} n. \qquad (1.26)$$

Аналогично, плотность дырочной составляющей диффузионного тока

$$\mathbf{J}_{p\,\text{диф}} = -\,qD_p \operatorname{grad} p. \qquad (1.27)$$

Заряд дырок положителен, поэтому направление вектора плотности диффузионного тока дырок должно совпадать с направлением их диффузии, т. е. должно быть противоположным направлению вектора градиента концентрации дырок. Следовательно, в правой части соотношения (1.27) должен сохраниться знак минус.

Одновременно с процессом диффузии неравновесных носителей происходит процесс их рекомбинации. Поэтому избыточная концентрация уменьшается в направлении от места источника этой избыточной концентрации носителей. Расстояние, на котором при одномерной диффузии в полупроводнике без электрического поля в нем избыточная концентрация носителей заряда уменьшается вследствие рекомбинации в $e = 2{,}718$ ... раза, называют *диффузионной длиной* (L). Иначе говоря, это расстояние, на которое носитель диффундирует за время жизни. Таким образом, диффузионная длина связана с временем жизни носителей соотношениями

$$L_n = \sqrt{D_n\tau_n}\,; \quad L_p = \sqrt{D_p\tau_p}\,, \tag{1.28}$$

где D_n и D_p — коэффициенты диффузии электронов и дырок.

В свою очередь, коэффициенты диффузии связаны с подвижностью носителей заряда соотношениями Эйнштейна

$$D_n/\mu_n = kT/q; \quad D_p/\mu_p = kT/q. \tag{1.29}$$

Не следует путать диффузионную длину с *длиной свободного пробега* носителей заряда, которая определяется как среднее расстояние, проходимое носителем между двумя последовательными актами рассеяния.

Необходимо отметить, что диффузия носителей заряда может происходить в полупроводнике, имеющем первоначально равномерное распределение концентрации носителей, т. е. равный нулю градиент концентрации, но при наличии в полупроводнике разности температур или градиента температуры. В этом случае носители заряда (например, электроны), находящиеся в местах полупроводника с большей температурой, будут иметь бóльшую энергию, т. е. будут занимать более высокие энергетические уровни в зоне проводимости. В местах полупроводника с меньшей температурой энергетические уровни с аналогичной энергией свободны от электронов. Поэтому возникает диффузия электронов из нагретых мест полупроводника в холодные места. Результатом такой диффузии является возникновение градиента концентрации носителей заряда. Это явление аналогично процессу диффузии в газе — при нагревании давление повышается и частицы газа диффундируют в области с низкой температурой. Процессы, происходящие в полупроводниковых приборах, часто рассматривают, пренебрегая разностью температур между различными областями полупроводниковой структуры прибора. В этом случае для расчета диффузионных токов можно использовать соотношения (1.26) и (1.27).

Уравнения токов

При наличии электрического поля и градиента концентрации носителей заряда в полупроводнике будут существовать дрейфовые и диффузионные токи. Поэтому плотность электронного тока с учетом (1.22) и (1.26)

$$\mathbf{J}_n = \mathbf{J}_{n\,\text{др}} + \mathbf{J}_{n\,\text{диф}} = qn\mu_n\mathbf{E} + qD_n\,\text{grad}\,n. \tag{1.30}$$

Плотность дырочного тока с учетом (1.23) и (1.27)

$$\mathbf{J}_p = \mathbf{J}_{p\,\text{др}} + \mathbf{J}_{p\,\text{диф}} = qp\mu_p\mathbf{E} - qD_p\,\text{grad}\,p. \tag{1.31}$$

Для расчета плотности полного тока следует сложить его электронную и дырочную составляющие и добавить к ним плотность тока смещения, равную производной по времени от вектора электрической индукции:

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p + \frac{\partial}{\partial t}(\varepsilon\varepsilon_0\mathbf{E}). \tag{1.32}$$

§ 1.8. ТЕМПЕРАТУРНЫЕ ЗАВИСИМОСТИ КОНЦЕНТРАЦИИ НОСИТЕЛЕЙ ЗАРЯДА И ПОЛОЖЕНИЯ УРОВНЯ ФЕРМИ

Температурные зависимости концентрации носителей заряда и положения уровня Ферми в широком диапазоне температур и при различных концентрациях примесей представлены на рис. 1.9.

Рассмотрим характер кривой, соответствующей относительно малой концентрации примесей (доноров) $N_д$. В области низких температур с увеличением температуры и, следовательно, с увеличением энергии теплового движения концентрация свободных электронов растет за счет ионизации доноров (участок кривой между точками *1* и *2*). Угол наклона этого участка кривой характеризует энергию ионизации примесей. В данном диапазоне температур уровень Ферми находится между дном зоны проводимости и энергетическими донорными уровнями. При некоторой температуре, которой соответствует точка *2* кривой, вероятность заполнения донорных уровней оказывается равной 50%. Поэтому уровень Ферми при температуре T_2 совпадает с энергетическим уровнем донора.

При дальнейшем увеличении температуры концентрация свободных электронов практически не увеличивается (участок кривой между точками *2* и *3*), так как все примеси уже ионизированы, а вероятность ионизации собственных атомов полупроводника еще ничтожно мала. Участок кривой, соответствующий постоянной концентрации носителей заряда, называют участком истощения примесей. Первые два участка кривой (*1—2* и *2—3*) соответствуют примесной электропроводности полупроводника.

При относительно больших температурах (участок кривой за точкой *3*) концентрация свободных электронов (носителей заряда) растет с увеличением температуры вследствие перехода электронов через запрещенную зону. Наклон этого участка кривой характеризует ширину запрещенной зоны $\Delta Э$ полупроводника. Уровень Ферми при этих температурах расположен вблизи середины запрещенной зоны, а полупроводник можно считать собственным, так как концентрация носителей заряда определяется ионизацией собственных атомов полу-

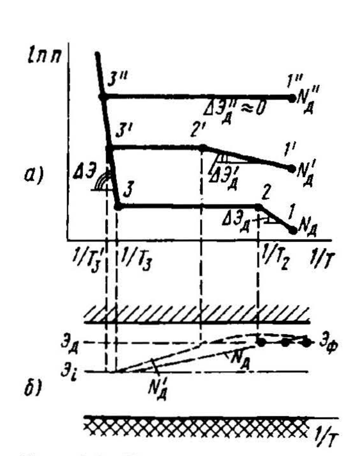


Рис. 1.9. Температурные зависимости концентрации свободных электронов в полупроводнике при различных концентрациях доноров $N_д < N_д' < N_д''$ (*а*) и соответствующие зависимости положения уровня Ферми (*б*)

проводника. Температура, при которой наступает собственная электропроводность или при которой полупроводник становится собственным, тем меньше, чем меньше ширина запрещенной зоны полупроводника. Принцип действия большинства полупроводниковых приборов основан на использовании примесной электропроводности, и поэтому появление собственной электропроводности нарушает нормальную работу прибора. Таким образом, температура, которой соответствует точка *3* кривой, является максимальной рабочей температурой полупроводникового прибора, изготовленного из полупроводника с концентрацией примеси $N_д$.

Теперь рассмотрнм смещение кривых и некоторое измененеие их характера при увелнчении концентрации примесей ($N''_д > N'_д > N_д$).

С увеличением концентрации примесей участкн кривых, соответствующие примесной электропроводности, смещаются вверх, т. е. получается большая концентрация носителей заряда при температурах примесной электропроводности.

Угол наклона первого участка кривой (участка ионизации примесей) с увеличением концентрации примесей уменьшается, так как с увеличением концентрации примесей из-за взаимодействия примесных атомов происходит расщепление примесных энергетических уровней и уменьшение энергии ионизацин примесей. Поэтому $\Delta Э_д > \Delta Э'_д > \Delta Э''_д$. При достаточно большой концентрации примесей ($N''_д$) энергия ионизации примесей стремится к нулю, так как образовавшаяся примесная зона перекрывается с зоной проводимости. Такой полупроводник является вырожденным (полуметаллом).

Температура, соответствующая переходу от примесной электропроводности к собственной, увеличивается с увеличением концентрации прнмесей (например, $T'_3 > T_3$). Это значит, что максимальная рабочая температура полупроводникового прибора, созданного на основе полупроводника с большей концентрацией примесей, будет также немного выше максимальной рабочей температуры такого же прибора из того же материала, но с меньшей концентрацией примесей.

§ 1.9. ТЕМПЕРАТУРНЫЕ ЗАВИСИМОСТИ ПОДВИЖНОСТИ НОСИТЕЛЕЙ ЗАРЯДА И УДЕЛЬНОЙ ПРОВОДИМОСТИ

На подвижность носителей заряда в основном влияют два физических фактора: хаотические тепловые колебания атомов кристаллической решетки (рассеяние носителей заряда на тепловых колебаниях атомов кристаллической решетки) и электрические поля ионизированных примесей (рассеяние на ионах примесей). При больших температурах преобладает рассеяние носителей заряда на тепловых колебаниях атомов кристаллической решетки.

Поэтому с увеличением температуры в этом диапазоне температур подвижность носителей уменьшается (рис. 1.10).

В диапазоне малых температур с уменьшением температуры уменьшаются тепловые скорости хаотического движения носителей заряда, что приводит к увеличению времени пребывания носителя вблизи иона примеси, т. е. увеличивается длительность воздействия электрического поля иона примеси на носитель заряда. Поэтому в диапазоне малых температур с уменьшением температуры подвижность носителей также уменьшается (рис. 1.10).

При увеличении концентрации примесей увеличивается и рассеяние на ионах примесей, т. е. уменьшается подвижность носителей заряда. Однако в диапазоне высоких температур преобла-

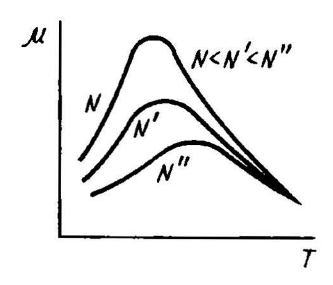


Рис. 1.10. Температурные зависимости подвижности носителей заряда при различных концентрациях примесей

Рис. 1.11. Температурные зависимости удельной проводимости полупроводника при различных концентрациях примесей

дающим механизмом рассеяния носителей даже при большой концентрации примесей остается рассеяние на тепловых колебаниях атомов кристаллической решетки и соответственно кривые температурной зависимости подвижности носителей заряда в диапазоне высоких температур практнчески не смещаются с увеличением концентрации примесей.

Удельная проводимость пропорцнональна концентрации носителей заряда и их подвижности. Поэтому, зная влияние температуры на концентрацию и подвижность носителей заряда, можно представить и общий ход кривой, отражающей зависимость удельной проводимости от температуры (рис. 1.11). Концентрация носителей заряда в полупроводниках очень сильно зависит от температуры — по экспоненциальному закону (1.7), а на подвижность изменение температуры влняет сравнительно слабо — по степеннóму закону (исключение составляют оксидные полупроводники на основе оксидов металлов с перемеиной валентностью). Поэтому температурная зависимость удельной проводимости похожа на температурную зависимость концентрации носителей прн очень малых и при больших температурах. В диа-

пазоне температур, соответствующих истощению примесей, когда концентрация основных носителей заряда остается практически неизменной, температурные изменения удельной проводимости обусловлены температурной зависимостью подвижности.

§ 1.10. ПОЛУПРОВОДНИКИ В СИЛЬНЫХ ЭЛЕКТРИЧЕСКИХ ПОЛЯХ

В сильных электрических полях в полупроводнике могут происходить физические процессы, приводящие к изменению удельной проводимости полупроводника; вольт-амперная характеристика полупроводника перестает подчиняться закону Ома; может изменяться как концентрация носителей заряда, так и их подвижность.

Рассмотрим вначале физические процессы, влияющие на концентрацию носителей заряда.

Ударная ионизация

Свободный электрон (или дырка), разгоняясь под действием большой напряженности электрического поля, может приобрести на длине свободного пробега дополнительную энергию, достаточную для ионизации примеси или собственного атома полупроводника. Процесс ионизации атомов разогнавшимся в поле носителем заряда называют *ударной ионизацией*. Ионизацию могут вызывать и дырки, так как движение дырок является лишь способом описания движения совокупности электронов валентной зоны полупроводника.

Количественно процесс ударной ионизации характеризуется *коэффициентами ударной ионизации*, которые численно равны количеству пар носителей заряда, образуемых первичным носителем на единице пути. По аналогии с теорией электрического разряда в газах, коэффициенты ударной ионизации в полупроводниках обозначают α_n и α_p. Коэффициенты ударной ионизации очень сильно зависят от напряженности электрического поля. Для практических расчетов часто пользуются эмпирической аппроксимацией

$$\alpha = A\,|E|^m, \qquad (1.33)$$

где m — довольно большой показатель степени, различный для разных материалов (от 5 до 8).

Туннелирование

Сильному электрическому полю в полупроводнике соответствует большой наклон энергетических зон (рис. 1.12). При этом электроны могут проходить сквозь узкий потенциальный барьер (толщиной Δ) без изменения своей энергии — *туннелировать* благо-

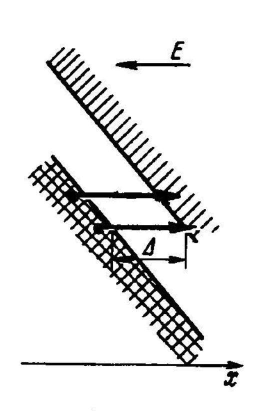


Рис. 1.12. Туннелироваине электронов из валентной зоны в зону проводимости при сильном электрическом поле в полупроводнике

даря своим квантово-механическим свойствам. Так как процесс туннелирования происходит вследствие перехода электронов из валентной зоны в зону проводимости, то этот процесс можио считать аналогичным автоэлектронной эмиссии или холодной эмиссии электронов из металла.

Вероятность перехода электронов из валентной зоны в зону проводимости и, наоборот, из зоны проводимости в валеитную зону одна и та же. Но переход электроиов из валентной зоны преобладает, поскольку их там значнтельно больше, чем в зоие проводимости. Поэтому концентрация носителей заряда растет при туннелировании.

Туннельный эффект в полупроводниках проявляется при очень больших иапряженностях электрического поля: в кремнии — при $E \approx 10^6$ В/см, в германии — при $E \approx \approx 10^5$ В/см. Напряжениости электрического поля, при которых появляется эффект туннелирования, различны для разных материалов, так как толщина потенциального барьера (Δ) зависит от ширнны запрещенной зоны полупроводника при неизменной напряженности электрического поля, т. е. при неизменном наклоне энергетических зон.

Теперь рассмотрим влняние сильного электрического поля на подвнжность носителей заряда.

Рассеяние носителей заряда в сильных полях

В слабых электрическнх полях носители заряда иа длине свободиого пробега приобретают относительно малую энергию. Поэтому их распределение по энергетическим уровням соответствует распределению при данной температуре кристаллической решетки. Дрейфовые скорости двнжения носителей заряда при этом значительно меньше так называемых тепловых скоростей, т. е. скоростей тепловых хаотическнх движеиий.

В снльных электрических полях скорость дрейфа носителей заряда соизмерима с тепловой скоростью; носители заряда на длине свободного пробега приобретают в электрическом поле энергии, соответствующне кинетическим энергиям теплового хаотического движения. При этом распределение носителей заряда по энергетическнм уровням соответствует бóльшим температурам, чем температура кристаллической решетки, которая остается практнчески неизменной. Это явление называют иногда *разогревом носителей*. На подвижность носителей явление разогрева может влиять по-разному.

1. При относительно больших температурах, при которых подвижность носителей заряда определяется в основном процес-

сом рассеяния на тепловых колебаннях атомов кристаллической решетки полупроводника, разогрев носителей заряда электрнческим полем приводит к увеличению числа столкновений носителей с атомами кристаллической решетки, т. е. к насыщению дрейфовой скорости или к уменьшению подвижности при увеличении напряженности электрического поля (рис. 1.13). Именно это явление надо учитывать в полупроводнниковых прнборах, еслн электрические поля превышают значение 10^3—10^4 В/см.

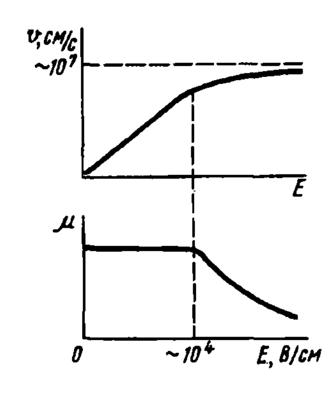


Рис. 1.13. Зависимость дрейфовой скорости и подвижности носителей заряда от напряженности электрического поля

2. При относительно малых температурах, при которых подвижность носителей заряда определяется в основном процессом рассеяния на иоиизированных примесях, разогрев носителей электрическим полем приводит к уменьшению времени нахождения носителя в поле ионизироваиной примеси, т. е. к уменьшению рассеяння носителя и, следовательно, к увеличению подвижности. Таким образом, увеличение подвижности с увеличением напряженности электрнческого поля в полупроводниковых приборах может происходить только при очень низких температурах.

Междолиииый переход носителей заряда

Рассмотрим вначале зависимость энергни свободного электрона, находящегося в вакууме, от его импульса **P** (рис. 1.14). Энергия такого электрона

$$Э = \frac{m_0 v^2}{2} = \frac{m_0^2 v^2}{2m_0} = \frac{P^2}{2m_0},$$

где **v** — вектор скорости свободного электрона; m_0 — его масса.

Зависимость, представленная на рис. 1.14, является энергетической диаграммой свободных электронов в вакууме, изображенной в пространстве импульсов или в пространстве волновых векторов **k** $\left(\mathbf{P} = \hbar \mathbf{k} = \frac{h}{2\pi}\mathbf{k}\right)$.

В полупроводниковом крнсталле свободный электрон можно считать свободным только условно, так как на электрон в кристалле действует периодическое потенциальное поле кристаллической решетки. Чтобы описать сложные законы движения электрона в кристалле с помощью соотношений, совпадающих по форме с законами классической механики, можно учесть влияние внутренних сил на электрон, изменив соответствующим образом значение его массы, т. е. введя понятие некоторой эф-

фективной массы электрона (или дырки). Таким образом, *эффективная масса* — это коэффициент пропорциональности в законе, связывающем внешнюю силу, действующую на электрон в кристалле, с его ускорением.

Зона проводимости полупроводника может быть образована из нескольких перекрывающихся между собой разрешенных энергетических зон. Структура энергетических зон или энергетическая диаграмма полупроводника в пространстве квазиимпульсов (в k-пространстве) может иметь несколько минимумов (рис. 1.15). Например, на энергетической диаграмме арсенида

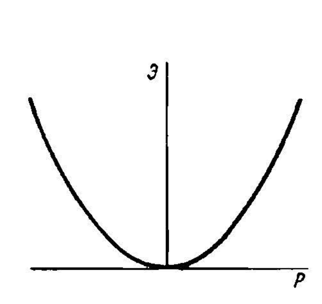


Рис. 1.14. Зависимость энергии свободного электрона, находящегося в вакууме, от его импульса

Рис. 1.15. Структура энергетических зон арсенида галлия в кристаллографическом направлении [100]

галлия зона проводимости кроме центральной долины с минимумом энергии при волновом векторе $\mathbf{k} = 0$ имеет еще боковые долины с минимумом энергии, который отличается от предыдущего на $\Delta Э_1$.

Из этой энергетической диаграммы следует, что в зоне проводимости арсенида галлия могут существовать электроны, обладающие одной и той же энергией, но имеющие различные квазиимпульсы, а следовательно, и различные эффективные массы:

$$Э = \frac{\mathbf{P}_1^2}{2m_1^*} = \frac{\mathbf{P}_2^2}{2m_2^*}. \qquad (1.34)$$

Если $\mathbf{P}_1 < \mathbf{P}_2$, то $m_1^* < m_2^*$, а это значит, что $\mu_1 > \mu_2$. Таким образом, в полупроводнике (арсениде галлия) могут существовать свободные электроны с разными подвижностями: «легкие» электроны с малой эффективной массой и с большой подвижностью в центральной долине и «тяжелые» электроны с большой эффективной массой и с малой подвижностью в боковых долинах.

В слабых электрических полях почти все свободные электроны имеют малые дрейфовые скорости и квазиимпульсы и поэто-

му находятся в центральной долине. В сильных электрических полях свободные электроны, приобретая дополнительную энергию, превышающую $\Delta Э_1$, получают возможность перейти в боковые долины. Там они характеризуются большей эффективной массой (станут «тяжелыми») и малой подвижностью. По этой причине средняя подвижность всех свободных электронов с увеличением напряженности электрического поля уменьшается. Подвижности «легких» и «тяжелых» электронов могут отличаться в десятки раз.

§ 1.11. Оптические свойства полупроводников

Поглощение света

Существуют различные виды поглощения света. При поглощении полупроводником квантов света — фотонов — их энергия может быть передана электронам валентной зоны с переводом этих электронов в зону проводимости, т. е. энергия квантов света идет на ионизацию атомов полупроводника. Этот процесс называют *собственным поглощением*. Существует поглощение энергии квантов света свободными электронами зоны проводимости или дырками валентной зоны, т. е. *поглощение носителями заряда*. При этом энергия квантов света расходуется также на перевод носителей на более высокие для них энергетические уровни, но в пределах соответствующей разрешенной зоны. Возможно *примесное поглощение*, при котором энергия фотонов идет на ионизацию или возбуждение примесных атомов. Кроме того, в полупроводниках могут происходить поглощение фотонов кристаллической решеткой и некоторые другие виды поглощения. Процессы поглощения фотонов не следует смешивать с процессами рассеяния, которые также приводят к уменьшению плотности потока фотонов.

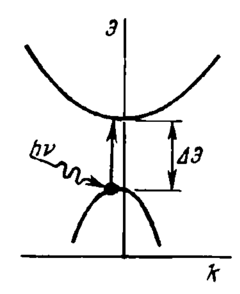


Рис. 1.16. Прямой переход электрона из валентной зоны в зону проводимости полупроводника

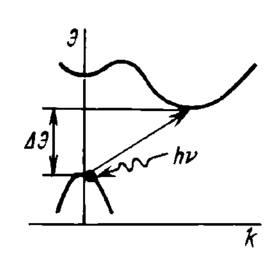


Рис. 1.17. Непрямой переход электрона из валентной зоны в зону проводимости полупроводника

При собственном поглощении фотонов переход электронов из валентной зоны в зону проводимости полупроводника может происходить без изменения квазиимпульса или волнового вектора электрона, т. е. возможны *прямые переходы* (рис. 1.16). Может происходить также переброс электронов из валентной зоны в зону проводимости и с изменением волнового вектора — *непрямые переходы* (рис. 1.17). При непрямых переходах в процессе поглощения кроме фотона и электрона должна участвовать еще третья квазичастица, которая заберет часть квазиимпульса на себя, т. е. обеспечит выполнение закона сохранения импульса. Таким образом, непрямые переходы — это переходы с участием третьей квазичастицы. Третьей квазичастицей обычно является фонон — квант тепловой энергии кристаллической решетки полупроводника.

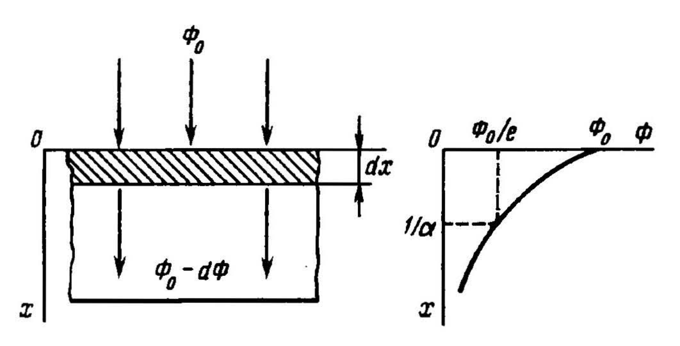


Рис. 1.18. Поглощение света в полупроводнике

Поглощение света или вообще фотонов характеризуют *показателем поглощения* α, который равен относительному изменению светового потока (потока фотонов) в слое полупроводника единичной толщины (рис. 1.18):

$$\alpha = -\frac{1}{\Phi}\frac{d\Phi}{dx}.$$

Это соотношение представляет собой дифференциальное уравнение с разделяющимися переменными. Поэтому

$$\int_{\Phi_0}^{\Phi}\frac{d\Phi}{\Phi} = -\alpha\int_{0}^{x}dx;\ \ln\frac{\Phi}{\Phi_0} = -\alpha x \text{ и } \Phi = \Phi_0\exp(-\alpha x).$$

Таким образом, показатель поглощения α можно определить как величину, обратную толщине слоя полупроводника, после прохождения которого световой поток (поток фотонов) уменьшится в $e = 2{,}718\ldots$ раза.

Зависимость показателя поглощения от энергии фотонов называют спектром поглощения полупроводника (рис. 1.19). При больших энергиях фотонов происходит собственное поглощение с образованием пар носителей электрон — дырка. Показатель поглощения при этом велик. При малой энергии фото-

нов (меньше ширины запрещенной зоны полупроводника) показатель поглощения уменьшается.

При еще меньших энергиях квантов света может происходить примесное поглощение, если не все примеси ионизированы при даниой температуре. Примесному поглощению соответствует один или несколько максимумов в спектре поглощения при энергиях квантов света, равных энергиям ионизации примесей.

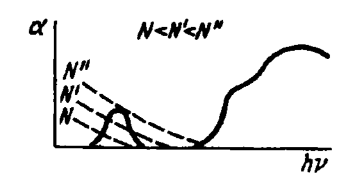


Рис. 1.19. Спектр поглощения полупроводника

При малых энергиях фотонов основным процессом поглощения является поглощеиие носителями заряда. Поэтому показатель поглощения в этом диапазоне частот зависит от концентрации носителей или от концентрации примесей.

Люминесценция полупроводинков

Люминесценцией называют электромагнитное нетепловое излучение, обладающее длительностью, значительно превышающей период световых колебаний.

Для возникновения люминесценции в полупроводнике атомы полупроводника должны быть выведены из состояния термодинамического равновесия, т. е. возбуждены. Они могут быть переведены в возбужденное состояние электрическим полем (*электролюминесценция*), бомбардировкой полупроводника электронами (*катодолюминесценция*), освещением (*фотолюминесценция*) и с помощью других энергетическнх воздействий. При люминесценции акты поглощения энергии полупроводником и излучения квантов света разделены во времени (а может быть, и в пространстве) промежуточными процессами, что приводит к относительно длительному существованию свечения полупроводника после прекращения возбуждения.

Излучение квантов света из полупроводника может происходить в результате перехода электрона на более низкий энергетический уровень при межзонной рекомбинации или при рекомбинации с участием рекомбинационных ловушек (см. рис. 1.3).

Излучательная рекомбинация носителей заряда может произойти без электромагиитного воздействия, т. е. самопроизвольно. Такую рекомбинацию называют *самопроизвольной* или *спонтанной*. Так как спонтанное излучение случайно и имеет статистический характер, то оно некогерентно. Это обусловлено тем, что акты спонтанного излучения происходят независимо друг от друга в разные моменты времени.

Переход электрона на более низкий энергетический уровень с излучением кванта света может произойти с помощью электромагнитного воздействия. Такую рекомбинацию (такой переход) называют *вынужденной* или *стимулированной*. Следовательно,

квант света с определенной частотой может не только поглощаться полупроводником, но и вызывать добавочное индуцированное излучение. Индуцированное излучение происходит в том же направлении, что и вызвавшие его излучение, в одной и той же фазе и с одинаковой поляризацией, т. е. индуцированное излучение является когерентным.

§ 1.12. ФОТОЭЛЕКТРИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

Фоторезистивный эффект

Фоторезистивный эффект — это изменение электрического сопротивления полупроводника, обусловленное исключительно действием оптического излучения и не связанное с его нагреванием. Для возникновения фоторезистивного эффекта необходимо, чтобы в полупроводнике происходило либо собственное поглощение оптического излучения или фотонов с образованием новых пар носителей заряда, либо примесное поглощение с образованием носителей одного знака при возбуждении однотипных дефектов. В результате увеличения концентрации носителей заряда уменьшается сопротивление полупроводника.

При облучении полупроводника наряду с генерацией неравновесных носителей заряда происходит и обратный процесс — рекомбинация. Через некоторое время после начала облучения устанавливается динамическое равновесие между генерацией и рекомбинацией. При этом избыточная концентрация, например, электронов

$$\Delta n = (1 - R)\alpha\eta N_{\text{ф}}\tau_n,$$

где R — коэффициент отражения фотонов от полупроводника; α — показатель поглощения; η — *квантовая эффективность генерации,* т. е. число возникающих пар носителей при собственном поглощении (или число носителей при примесном поглощении), отнесенное к числу поглощенных фотонов; $N_{\text{ф}}$ — число фотонов, падающих на единичную поверхность полупроводника в единицу времени (оно может быть определено как мощность падающего на единичную поверхность излучения, отнесенное к энергии фотона $h\nu$); τ_n — время жизни неравновесных носителей заряда.

Возникновение ЭДС в однородном полупроводнике при его освещении

Если однородный полупроводник осветить сильно поглощаемым светом, то в его поверхностном слое, где происходит основное поглощение света, возникнет избыточная концентрация электронов и дырок, которые будут диффундировать в глубь

полупроводника (рис. 1.20). Коэффициент диффузии электронов обычно значительно больше коэффициента диффузии дырок. Поэтому при диффузии электроны опережают дырки, происходит некоторое разделение зарядов — поверхность полупроводника приобретает положительный заряд, а объем заряжается отрицательно. Таким образом, в полупроводнике при его освещении возникает электрическое поле или ЭДС, которую иногда называют *ЭДС Дембера*. Возникшее электрическое поле будет тормозить электроны и ускорять дырки при их движении от поверхности полупроводника, в результате чего через некоторое время после начала освещения установится динамическое равновесие.

Напряженность электрического поля, возникающую в полупроводнике при его освещении, можно найти, используя уравнения (1.30)—(1.32), с учетом того, что в установившемся состоянии динамического равновесия тока через полупроводник нет. Тогда

$$0 = qn\mu_n E + qD_n \operatorname{grad} n + qp\mu_p E - qD_p \operatorname{grad} p.$$

Если же считать, что $\operatorname{grad} n \approx \operatorname{grad} p$, то

$$E = -\frac{D_n - D_p}{n\mu_n + p\mu_p} \operatorname{grad} p, \qquad (1.35)$$

Рис. 1.20. Возникновение ЭДС в полупроводнике при его освещении

т. е. напряженность электрического поля пропорциональна возникающему при освещении полупроводника градиенту концентрации носителей заряда.

§ 1.13. ОБЕДНЕННЫЕ, ИНВЕРСНЫЕ И ОБОГАЩЕННЫЕ ПОВЕРХНОСТНЫЕ СЛОИ

Поверхность полупроводника представляет собой нарушение периодичности кристаллической решетки. Из-за этого возникают дополнительные энергетические уровни, расположенные в запрещенной зоне энергетической диаграммы полупроводника. Эти уровни, теоретически предсказанные И. Е. Таммом, называют уровнями Тамма. Уровни Тамма являются акцепторными, так как у атомов полупроводника (например, кремния), находящихся у поверхности кристалла, оказывается всего по три соседних атома вместо четырех и, следовательно, отсутствует одна электронная связь. Плотность поверхностных уровней Тамма или поверхностных состояний должна быть того же порядка, что и число атомов на единице поверхности кристалла, т. е. около 10^{15} см$^{-2}$.

На реальной поверхности полупроводника поверхностные состояния возникают также вследствие адсорбции различных

примесей (кислорода, воды и других атомов, иоиов и молекул). Очевидно, что такая сложная структура поверхности реального полупроводника характеризуется дополнительными энергетическими уровнями в запрещенной зоне доиорного, акцепторного типа или типа ловушек.

Локальные энергетические уровни, обусловлеиные нарушением периодичности кристалла у поверхности полупроводника или примесями на поверхности, называют *поверхностными уровнями.*

Прн комнатной температуре большинство примесей обычно ионизироваио, т. е. на поверхностиых уровнях находятся заряды. Для компеисации этих зарядов в соответствии с условием

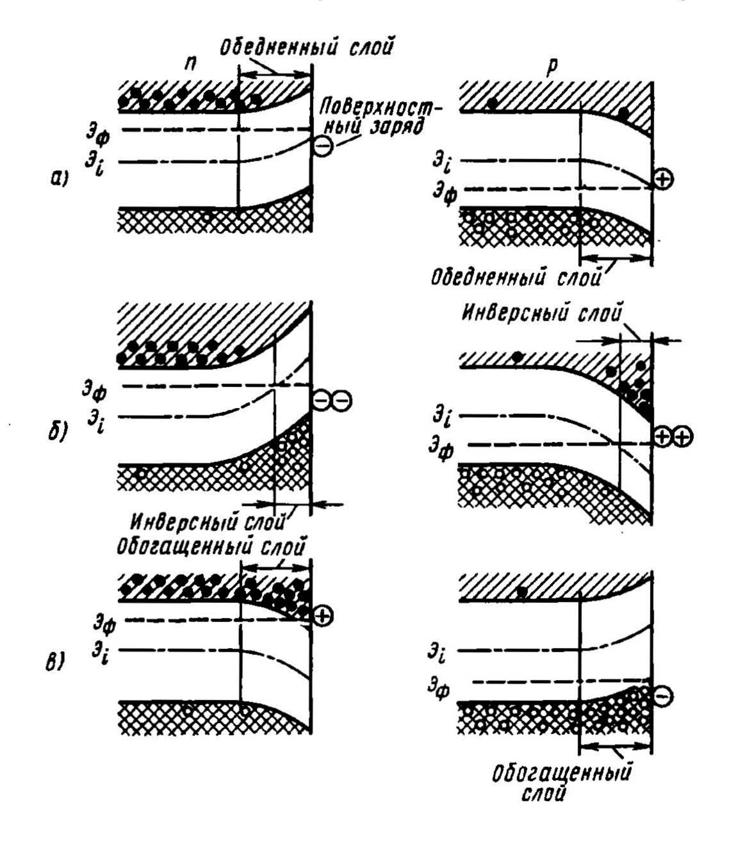


Рис. 1.21. Искажение энергетических зон у поверхности полупроводника н образование различных поверхностных слоев

электрической нейтральиости должеи существовать объемный заряд в полупроводнике, что соответствует существованию электрического поля и изгибу энергетических зои вблизи поверхиости полупроводника.

На рис. 1.21, *а—в* показано образование трех возможных вариантов поверхностных слоев в полупроводниках *n*- и *p*-типа при наличии положительных или отрицательных поверхностных состояний.

При малой плотности отрицательных поверхностных состояний на полупроводнике *n*-типа образуется *обедненный слой* (рис. 1.21, *а*), так как основные носители заряда — электроны — отталкиваются отрицательным поверхностным зарядом в глубь полупроводника. Электростатическое поле поверхностных зарядов проникает в полупроводник на определенную глубину, которая зависит от удельного сопротивления этого полупроводника.

При большой плотности отрицательных поверхностных состояний у поверхности полупроводника *n*-типа образуется слой с противоположным типом электропроводности — *инверсный слой*. Граница инверсного слоя в глубине полупроводника расположена там, где уровень Ферми пересекает середину запрещенной зоны. Под инверсным слоем в полупроводнике находится обедненный слой.

Если на поверхности полупроводника *n*-типа преобладают положительные поверхностные состояния, то поверхность полупроводника обогащается электронами — основными носителями заряда, т. е. образуется *обогащенный слой*. Граница обогащенного слоя в глубине полупроводника находится там, где начинается изгиб энергетических уровней, т. е. определяется глубиной проникновения электростатического поля поверхностных зарядов.

Таким образом, толщина области объемного заряда зависит от плотности поверхностных состояний, от удельного сопротивления полупроводника или от концентрации примеси и составляет обычно 10^{-2}—10 мкм.

На поверхности полупроводника *p*-типа могут образовываться аналогичные поверхностные слои, но при других знаках поверхностных зарядов (рис. 1.21).

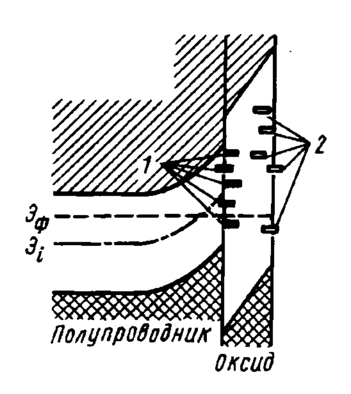


Рис. 1.22. Быстрые *1* и медленные *2* поверхностные состояния на реальной поверхности полупроводника

На реальном полупроводнике всегда имеется слой оксида. Поэтому поверхностные состояния могут находиться не только непосредственно на полупроводнике, но также в слое оксида и на его поверхности. При изменении внешнего электрического поля и при соответствующем изменении энергетической диаграммы вблизи поверхности полупроводника должно происходить заполнение или опустошение электронами по крайней мере некоторых поверхностных состояний. Поверхностные состояния, расположенные вблизи границы раздела

полупроводник — оксид, заполняются или опустошаются относительно быстро, так как в полупроводнике есть достаточное количество свободных электронов в зоне проводимости и дырок (пустых энергетических уровней) в валентной зоне. Из-за малого времени релаксации таких поверхностных состояний (10^{-8}—10^{-4} с) их называют *быстрыми поверхностными состояниями* (рис. 1.22).

В слое оксида концентрация носителей заряда ничтожно мала. Поэтому с изменением внешнего электрического поля поверхностные состояния, расположенные в слое оксида или на его поверхности, перезаряжаются медленно (время их релаксации велико). Такие поверхностные состояния называют *медленными поверхностными состояниями*. Время релаксации медленных поверхностных состояний составляет обычно от миллисекунд до нескольких часов.

§ 1.14. ПОВЕРХНОСТНАЯ РЕКОМБИНАЦИЯ

Некоторые поверхностные состояния создают вблизи середины запрещенной зоны энергетические уровни, которые являются уровнями рекомбинационных ловушек. Рекомбинацнонными ловушками могут быть только быстрые поверхностные состояния, так как время перехода носителей на медленные поверхностные состояния очень велико. Явление поверхностной рекомбинации принято характеризовать *скоростью поверхностной рекомбинации* носителей заряда, которая определяется как отношение плотности потока носителей заряда у поверхности полупроводника к избыточной концентрации этих носителей у поверхности, т. е.

$$s_p = \frac{\Phi_p}{\Delta p} = \frac{J_p}{q\Delta p}. \qquad (1.36)$$

Таким образом, скорость поверхностной рекомбинации показывает, какое количество носителей заряда рекомбинирует за 1 с на поверхности полупроводника площадью 1 см^2. Размерность скорости поверхностной рекомбинации та же, что и размерность скорости движения, т. е. см/с — скорость.

Плотность поверхностных состояний вообще и поверхностных состояний, являющихся рекомбинационными ловушками, в частности, зависит от обработки поверхности полупроводника и от свойств внешней среды, с которой он соприкасается. Так как плотность поверхностных состояний обычно велика, то рекомбинация носителей на поверхности идет значительно быстрее (интенсивнее), чем в объеме полупроводника. При малых размерах кристалла полупроводника явление поверхностной рекомбинации будет существенно уменьшать эффективное время жизни носителей заряда, так как

$$1/\tau_{эф} = 1/\tau_V + 1/\tau_s,$$

где $\tau_{эф}$ — эффективное время жизни; τ_V — время жизни в объеме полупроводника; τ_s — время жизни на поверхности полупроводника.

Кроме того, плотность поверхностиых состояний может изменяться со временем из-за испарения или конденсации влаги иа поверхностн кристалла, из-за возможных миграций адсорбированных примесей на поверхности и т. п. Эти процессы, приводя к изменению эффективного времеии жизни носителей заряда, могут являться причиной нестабильности параметров и характернстик полупроводниковых приборов.

Таким образом, при изготовлеиии полупроводииковых приборов необходимо, во-первых, выбирать метод обработки поверхности кристаллов полупроводника, при котором скорость поверхностиой рекомбинации миннмальиая, и, во-вторых, находить способ длительного сохранения достигнутых зиачений скорости поверхностной рекомбииации. Последнюю задачу обычно решают посредством нанесения иа поверхиость кристалла специальиых покрытнй н герметизации прибора в корпус.

§ 1.15. ПРОВОДИМОСТЬ КАНАЛА ПОВЕРХНОСТНОЙ ЭЛЕКТРОПРОВОДНОСТИ

При наличии на поверхности полупроводиика иверсиого слоя по иему вдоль поверхности могут проходить токи, т. е. могут существовать каиалы поверхиостной электропроводности. При оцеике проводимости каиала поверхиостной электропроводности необходимо учитывать, что концеитрация носителей заряда является величииой перемеиной по глубиие канала и подвижность носителей заряда в канале может значительно отличаться от подвижности тех же носителей в объеме полупроводника из-за дополиительного рассеяния носителей на поверхиостн кристалла.

Проводимость тоикого слоя (толщнной dz) канала, иапример, с электропроводностью p-типа

$$d\sigma = qp\mu_p dz\frac{b}{l},$$

где b — ширииа канала; l — длина канала.

Проводимость канала толщииой W

$$\sigma = q\frac{b}{l}\int_0^W p\mu_p dz,$$

его сопротивлеиие

$$R = \frac{1}{\sigma} = \frac{1}{q\int_0^W p\mu_p dz}\frac{l}{b}.$$

Вместо сопротивлеиия при оцеике токов в канале пользуются другим параметром — удельным сопротйвлеиием слоя (ϱ_s),

аналогичным удельному поверхностному сопротивлению диэлектриков, т. е. *удельным сопротивлением слоя* называют сопротивление квадрата этого слоя току, проходящему между двумя противоположными сторонами квадрата.

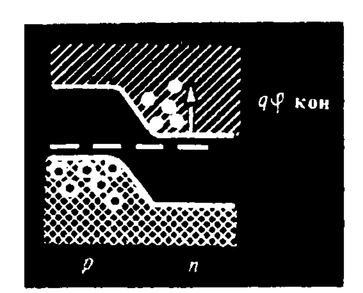
Тогда, используя удельное сопротивление слоя, запишем

$$R = \varrho_s \frac{l}{b},$$

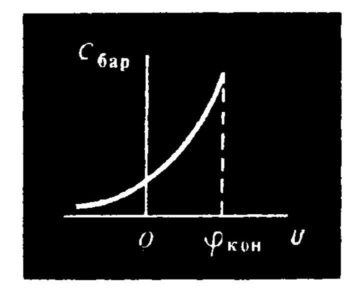
где $\varrho_s = \dfrac{1}{q\int_0^W p\mu_p dz}$ — удельное сопротивление слоя или канала поверхностной электропроводности, Ом.

Контрольные вопросы

1. Что такое разрешенные и запрещенные энергетические зоны?
2. Что такое ширина запрещенной зоны?
3. Каковы правила наиболее целесообразного построения энергетических диаграмм?
4. Чем определяется наклон (угловой коэффициент) энергетических уровней и зон на энергетической диаграмме полупроводника?
5. Что такое уровень Ферми?
6. В каких случаях и почему надо пользоваться различными статистиками распределения электронов по энергиям?
7. Что такое собственный полупроводник?
8. Чему равно произведение концентрации электронов и дырок в невырожденном полупроводнике при термодинамическом равновесии?
9. Что такое диффузия и дрейф носителей заряда?
10. Что такое подвижность носителей заряда?
11. Что такое диффузионная длина и длина свободного пробега носителей заряда?
12. Как объяснить температурную зависимость концентрации носителей заряда в полупроводнике?
13. Какими физическими факторами объясняется температурная зависимость подвижности носителей заряда?
14. Как и почему изменяется рассеяние носителей заряда в сильных электрических полях?
15. Что такое прямые и непрямые переходы носителей заряда между разрешенными зонами полупроводника?
16. Что такое показатель поглощения света полупроводником?
17. Что такое фоторезистивный эффект?
18. Какие разновидности поверхностных слоев могут возникать на полупроводнике при наличии различных поверхностных состояний?
19. Что такое скорость поверхностной рекомбинации?
20. Что такое удельное сопротивление слоя и какова размерность этого параметра?



2



Контактные явления

§ 2.1. ЭЛЕКТРОННО-ДЫРОЧНЫЙ ПЕРЕХОД

Электронно-дырочный переход (*p-n* - переход) — это переходный слой между двумя областями полупроводника с разной электропроводностью, в котором существует диффузионное электрическое поле.

Образование электронно-дырочного перехода

При идеальном контакте двух полупроводииков с различным типом электропроводности из-за градиента коицеитрации носителей заряда возиикает их диффузия в области с противоположным типом электропроводиости через плоскость *металлургического контакта* (плоскость, где измеияется тип примесей, преобладающих в полупроводнике). В результате диффузии носителей заряда нарушается электрическая иейтральиость примыкающих к металлургическому коитакту частей монокристалла полупроводиика. В *p*-области вблизи металлургического коитакта после диффузии из иее дырок остаются иескомпеисироваииые иоиизироваииые акцепторы (отрицательные неподвижные заряды), а в *n*-области — нескомпеисироваииые ионизированые доиоры (положительные иеподвижиые заряды). Образуется область простраиствеииого заряда, состоящая из двух разиоимеиио заряжеииых слоев. Между нескомпеисироваиными разиоимеиными зарядами ионизироваииых примесей возиикает электрическое поле, иаправлениое от *n*-области к *p*-области и называемое *диффузионным электрическим полем* (рис. 2.1, *а*). Возникшее диффузиоииое электрическое поле препятствует дальиейшей диффузии осиовных носителей через металлургический коитакт — устаиавливается равиовесное состояиие. Между *n*- и

p-областями при этом существует разность потенциалов, называемая *контактной разностью потенциалов*. Потенциал *n*-области положителен по отношению к потенциалу *p*-области.

Энергетическая диаграмма электронно-дырочного перехода

Согласно сказанному в § 1.1, энергетическую диаграмму электронно-дырочного перехода при термодинамическом равновесии можно изобразить, как показано на рис. 2.1, *г*. Вдали от контакта двух областей электрическое поле отсутствует (если соответствующие области легированы равномерно) или относительно мало по сравнению с полем в *p-n*-переходе. Поэтому энерге-

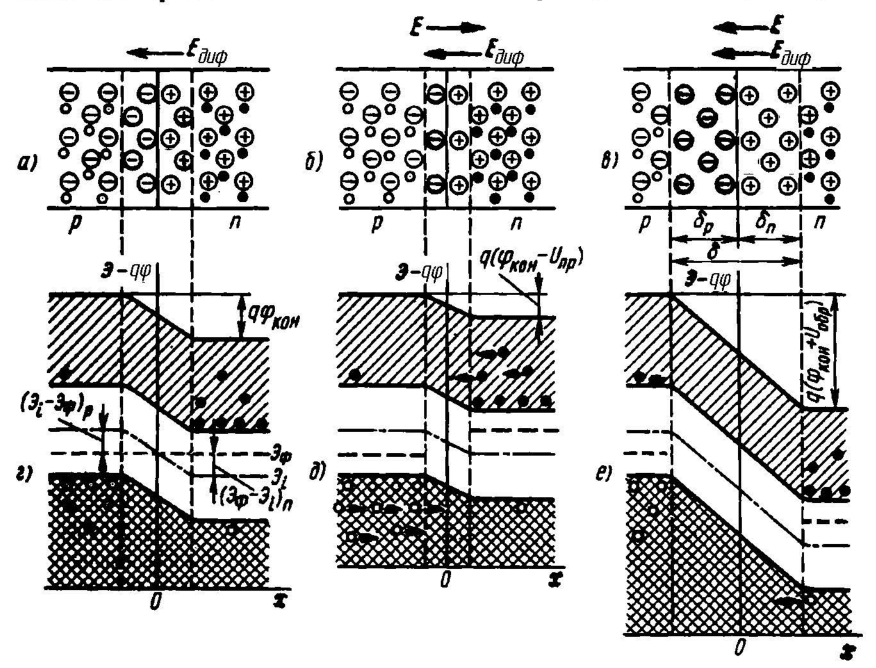


Рис. 2.1. Пространственное распределение зарядов (*а*, *б*, *в*) и энергетические диаграммы *p-n*-перехода (*г*, *д*, *е*,):

а, *г* — внешнее напряжение отсутствует ($U = 0$); *б*, *д* — внешнее напряжение прямое ($U > 0$); *в*, *е* — внешнее напряжение обратное ($U < 0$)

тические зоны в этих областях изображены горизонтальными. Взаимное расположение разрешенных зон и уровня Ферми за пределами *p-n*-перехода остается таким же, каким было в соответствующих полупроводниках.

Так как напряженность диффузионного электрического поля в *p-n*-переходе направлена от электронного полупроводника к

дырочному, на диаграмме соответствующие энергетические зоны для n-области должны быть ниже, чем для p-области. Сдвиг зон определяется тем, что в равновесном состоянии уровень Ферми должен быть расположен на одной высоте энергетической диаграммы всей системы. Этот сдвиг зон соответствует также контактной разности потенциалов ($\varphi_{кон}$) или высоте потенциального барьера ($q\varphi_{кон}$) электронно-дырочного перехода.

Высота потенциального барьера и контактная разность потенциалов

Как видно из рис. 2.1, *г*, высота потенциального барьера p-n-перехода

$$q\varphi_{кон} = (Э_{Ф} - Э_{i})_{n} + (Э_{i} - Э_{Ф})_{p}. \tag{2.1}$$

Пользуясь соотношениями (1.17) и (1.18), запишем

$$\ln\frac{n_{n0}}{n_i} = \frac{Э_{Ф} - Э_{i}}{kT} \text{ и } \ln\frac{p_{p0}}{n_i} = \frac{Э_{i} - Э_{Ф}}{kT}.$$

Следовательно, высота потенциального барьера

$$q\varphi_{кон} = kT\ln\frac{n_{n0}p_{p0}}{n_i^2} \tag{2.2}$$

или при использовании соотношения (1.19)

$$q\varphi_{кон} = kT\ln\frac{p_{p0}}{p_{n0}} = kT\ln\frac{n_{n0}}{n_{p0}}. \tag{2.3}$$

Все приведенные выражения для высоты потенциального барьера равносильны, но удобнее пользоваться формулой (2.2), так как концентрация основных носителей заряда практически равна концентрации легирующей примеси либо легко может быть определена из удельной проводимости прилегающих к p-n-переходу областей. Значение собственной концентрации носителей n_i для одной температуры дается в литературе, а температурная зависимость собственной концентрации определяется шириной запрещенной зоны полупроводника [см. (1.15)].

Для удобства анализа подставим (1.9) или (1.15) в (2.2) и учтем, что концентрация основных носителей заряда значительно меньше плотности возможных состояний в соответствующих зонах, т. е. $n_{n0}p_{p0} \ll 4N_{c}N_{в}$. Тогда

$$q\varphi_{кон} = \Delta Э - kT\ln\frac{4N_{c}N_{в}}{n_{n0}p_{p0}}. \tag{2.4}$$

Из соотношения (2.4) можно сделать следующие выводы:

1) при одних и тех же концентрациях примесей высота потенциального барьера больше в p-n-переходах, созданных в полупроводниках с большей шириной запрещенной зоны;

2) высота потенциального барьера возрастает при увеличении концентрации примесей в соответствующих областях;

3) с увеличением температуры высота потенциального барьера уменьшается.

§ 2.2. ТОКИ ЧЕРЕЗ ЭЛЕКТРОННО-ДЫРОЧНЫЙ ПЕРЕХОД

При нарушении равновесия внешним электрическим полем через *p-n*-переход начинает проходить ток. Если внешнее напряжение приложено так, что создаваемая им напряженность электрического поля противоположна по направлению диффузионной напряженности (рис. 2.1, *б*), то суммарная напряженность поля в *p-n*-переходе падает, высота потенциального барьера уменьшается (рис. 2.1, *д*). Часть основных носителей, имеющих наибольшие значения энергии, может теперь преодолевать понизившийся потенциальный барьер, переходя через *p-n*-переход. Это приводит к появлению сравнительно большого тока через *p-n*-переход. Напряжение рассмотренной полярности называют *прямым* и считают положительным.

Преодолевшие потенциальный барьер носители заряда оказываются в соседней области неосновными. Таким образом, через *p-n*-переход происходит *инжекция неосновных носителей* заряда в область, примыкающую к *p-n*-переходу. Ту область, в которую происходит инжекция неосновных носителей, называют *базой* полупроводникового прибора.

С увеличением внешнего прямого напряжения уменьшается суммарная напряженность электрического поля в *p-n*-переходе. С уменьшением напряженности электрического поля уменьшается глубина проникновения этого поля в области полупроводника, прилегающие к металлургическому контакту. Поэтому уменьшается толщина *p-n*-перехода или ширина области пространственного заряда.

Если созданное внешним источником электрическое поле в *p-n*-переходе совпадает по направлению с диффузионным (рис. 2.1, *в*), то высота потенциального барьера для основных носителей увеличивается (рис. 2.1, *е*). Однако для неосновных носителей, т. е. для дырок в *n*-области и для электронов в *p*-области, потенциальный барьер в *p-n*-переходе вообще отсутствует. Неосновные носители заряда втягиваются электрическим полем в *p-n*-переход и проходят через него в соседнюю область — происходит так называемая *экстракция*. При этом через *p-n*-переход будет идти обратный ток, который относительно мал из-за малой концентрации неосновных носителей заряда в прилегающих к *p-n*-переходу областях.

Напряжение, имеющее рассмотренную полярность, называют *обратным* и считают отрицательным. Толщина *p-n*-перехода с увеличением обратного напряжения по абсолютному значению увеличивается, так как при этом увеличивается суммарная

напряженность электрического поля в p-n-переходе и увеличивается глубина проникновеиия этого поля в прилегающие к переходу области.

§ 2.3. КОНЦЕНТРАЦИЯ НЕОСНОВНЫХ НОСИТЕЛЕЙ ЗАРЯДА У ГРАНИЦ ЭЛЕКТРОННО-ДЫРОЧНОГО ПЕРЕХОДА

Рассмотрим зависимость концентрации неосновных носителей заряда у границ p-n-перехода от внешнего напряжения, приложенного к электронно-дырочному переходу, для частных случаев.

Малые токи

Как известно, плотности электронного и дырочного токов определяются алгебраической суммой дрейфовых и диффузионных составляющих [см. (1.30) и (1.31)]. При движении носителей заряда только в одном направлении x, параллельном вектору электрического поля,

$$\mathbf{J}_p = -qp\mu_p \frac{d\varphi}{dx} - qD_p \frac{dp}{dx};$$

$$\mathbf{J}_n = -qn\mu_n \frac{d\varphi}{dx} + qD_n \frac{dn}{dx}.$$

Учитывая малость токов и воспользовавшись соотношением Эйнштейна (1.29), можно считать, что при $J_p \approx 0$ и $J_n \approx 0$

$$-\frac{q}{kT}\frac{d\varphi}{dx} = \frac{1}{p}\frac{dp}{dx} \text{ или } \frac{q}{kT}\int\limits_0^{\varphi_{кон}-u} d\varphi = -\int\limits_{p_p}^{p_{n,\,гр}} \frac{dp}{p};$$

$$\frac{q}{kT}\frac{d\varphi}{dx} = \frac{1}{n}\frac{dn}{dx} \text{ или } \frac{q}{kT}\int\limits_0^{\varphi_{кон}-u} d\varphi = \int\limits_{n_{p,\,гр}}^{n_n} \frac{dn}{n}.$$

При малых токах концентрации основных носителей за пределами p-n-перехода p_p и n_n практически равны равновесиым концентрациям p_{p0} и n_{n0}. Тогда с учетом (2.3)

$$p_{n,\,гр} = p_{n0} \exp\frac{qu}{kT};\quad n_{p,\,гр} = n_{p0} \exp\frac{qu}{kT}. \tag{2.5}$$

Полученные соотношения имеют довольно простой физический смысл. В невырожденном полупроводнике носители заряда подчиняются статистике Максвелла—Больцмана, т. е. число их с энергией выше некоторого значения экспоненциально падает с увеличением этой энергин. В состоянии равновесия концентра-

ция неосновных носителей заряда по одну сторону p-n-перехода равна концентрации основных носителей по другую сторону p-n-перехода, имеющих энергию, бóльшую на $q\varphi_{кон}$. При изменении высоты потенциального барьера на qu количество носителей, имеющих энергию, достаточную для его преодоления, изменяется в $\exp\frac{qu}{kT}$ раз, что и характеризуется формулами (2.5).

При напряжении на p-n-переходе, равном нулю, граничная концентрация неосновных носителей заряда равна равновесной. С ростом прямого напряжения ($u>0$) граничная концентрация неосновных носителей заряда растет, что соответствует явлению инжекции. При обратном напряжении ($u<0$) граничная концентрация неосновных носителей падает, что соответствует явлению экстракции.

Большие прямые токи

При увеличении прямого напряжения на p-n-переходе высота его потенциального барьера уменьшается, что ведет к выравниванию концентраций носителей по обе стороны от p-n-перехода. В пределе, если бы контактная разность потенциалов могла быть полностью скомпенсирована внешним напряжением (практически это недостижимо из-за падения части внешнего напряжения в объеме полупроводника), выполнялись бы условия

$$p_{n,\,\text{гр}\max}=p_p;\quad n_{p,\,\text{гр}\max}=n_n. \tag{2.6}$$

Соотношения (2.6) можно рассматривать как верхние пределы концентрации инжектированных через p-n-переход носителей.

Большие обратные токи

Соотношения (2.5) показывают, что концентрация неосновных носителей заряда на границе p-n-перехода, смещенного в обратном направлении, с ростом абсолютного значения обратного напряжения должна очень быстро падать. Однако такое падение концентрации ограничено тем, что скорость движения носителей заряда в электрическом поле растет до определенного предела $v_{\max}$ (см. § 1.10). Так как плотность тока, например дырочного, связана со скоростью движения носителей заряда соотношением

$$J_p=qpv_p, \tag{2.7}$$

то минимальное значение концентрации неосновных носителей заряда, которое может получиться на границе p-n-перехода,

$$p_{n,\,\text{гр}\min}=\frac{J_p}{qv_{p\max}};\quad n_{p,\,\text{гр}\min}=\frac{J_n}{qv_{n\max}}. \tag{2.8}$$

Выражения (2.8) можно рассматривать как нижние пределы концентрации неосновных носителей на границе p-n-перехода.

Зависимость граничной концентрации неосновных носителей заряда от напряжения

Чтобы подвести итог, построим в полулогарифмическом масштабе зависимость граничной концентрации неосновных носителей от напряжения на *p-n*-переходе (рис. 2.2). При малых токах эта

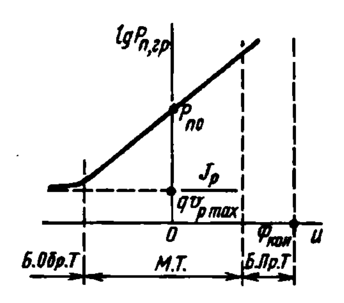


Рис. 2.2. Зависимость граничной концентрации неосновных носителей заряда около *p-n*-перехода от напряжения, приложенного к этому переходу:

Б. Обр. Т — диапазон больших обратных токов; М. Т. — диапазон малых токов; Б. Пр. Т — диапазон больших прямых токов

зависимость экспоненциальна, а в выбранном масштабе — прямая, проходящая через значение равновесной концентрации неосновных носителей заряда при $u=0$. При больших прямых токах, т. е. при напряжениях на *p-n*-переходе, стремящихся к контактной разности потенциалов, изображать зависимость не имеет смысла. При больших обратных токах граничная концентрация неосновиых носителей заряда стремится к значению $J_p/(qv_{p\,\max})$.

§ 2.4. МЕТОДЫ ФОРМИРОВАНИЯ И КЛАССИФИКАЦИЯ ЭЛЕКТРОННО-ДЫРОЧНЫХ ПЕРЕХОДОВ

Среди разнообразных методов формирования *p-n*-переходов наибольшее значение имеют два: *метод вплавления* и *метод диффузии* примесей. Электронно-дырочный переход, полученный методом вплавления в полупроводник (с последующей рекристалли-

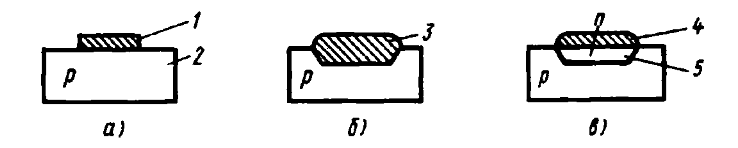


Рис. 2.3. Технологические стадии вплавления примесей в монокристалл полупроводника:

а — навеска вплавляемого материала *1*, содержащая донорную примесь, на поверхности монокристалла полупроводника *2* перед вплавлением; *б* — расплав материала навески с полупроводником *3* при высокой температуре вплавления; *в* — избыточный слой материала навески *4* на поверхности рекристаллизованного полупроводника *5* после охлаждения

зацией полупроводника) металла или сплава, содержащего донорные или акцепторные примеси, называют *сплавным* переходом, а переход, полученный в результате диффузии атомов примеси в полупроводник, — *диффузионным*.

При вплавлении в полупроводник металла или сплава, содержащего донорные или акцепторные примеси, полупроводник с навеской вплавляемого материала нагревают до расплавления навески, в результате чего часть кристалла полупроводника растворяется в расплаве навески. При последующем охлаждении происходит рекристаллизация полупроводникового кристалла с примесью вплавляемого материала. Если рекристаллизованный слой получился с другим типом электропроводности по сравнению с электропроводностью исходного полупроводника, то на границе их раздела возникает *p-n*-переход (рис. 2.3).

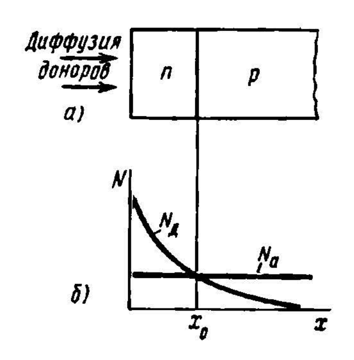


Рис. 2.4. Диффузионный *p-n*-переход (*а*) и распределение примесей в полупроводнике после диффузии в него доноров (*б*):

N_a — исходная концентрация акцепторов; $N_д$ — концентрация доноров после диффузии; x_0 — глубина залегания сформированного *p-n*-перехода

При создании диффузионного *p-n*-перехода используют диффузию в полупроводник примеси, находящейся в газообразной, жидкой или твердой фазе (рис. 2.4). Диффузионные *p-n*-переходы, в свою очередь, могут быть нескольких разновидностей. Так, диффузионный *p-n*-переход, образованный в результате диффузии примеси сквозь отверстие в защитном слое, нанесенном на поверхность полупроводника, называют *планарным p-n-переходом*. В качестве защитного слоя на кремнии обычно используют диоксид того же самого кремния.

На рис. 2.5 показаны этапы технологического процесса формирования планарного *p-n*-перехода. Основой такой технологии является фотолитография. На исходную окисленную пластину монокристалла кремния наносят слой фоточувствительного вещества — фоторезиста *ФР* (рис. 2.5, *а*). Пленку фоторезиста освещают через маску ультрафиолетовым светом (рис. 2.5, *б*). Экспонированные места фоторезиста полимеризуются и становятся нерастворимыми. После этого незаполимеризованные части фоторезиста смывают, так что он остается только на облученных местах (рис. 2.5, *в*). Затем производят травление пленки диоксида, которая остается только в тех местах, где она была защищена слоем фоторезиста (рис. 2.5, *г*). В дальнейшем проводят диффузию необходимой примеси в пластину исходного кремния. Диффузия примеси происходит селективно — только через отверстия или окна в слое диоксида кремния (рис. 2.5, *д*).

Для создания более сложных структур, например транзисторных с близко расположенными двумя p-n-переходами, необходимо еще раз повторить все рассмотренные этапы технологического процесса, т. е. заново окислить пластину кремния, нанести слой фоторезиста, засветить его определенные части поверхности, провести травление и диффузию примеси через вновь образованные окошки в слое диоксида кремния.

Электронно-дырочный переход, образованный в результате конверсии полупроводника, вызванной обратной диффузией примеси в соседнюю область, или активацией атомов примеси, называют *конверсионным p-n-переходом*. Так, для создания конверсионного p-n-перехода в германиевом монокристалле применяют германий, содержащий два типа примесей: донорную и акцепторную (медь). Концентрация меди в германии должна быть больше концентрации доноров, т. е. исходный германий имеет электропроводность p-типа. В такой германий проводят вплавление навески металла или сплава. Медь, отличаясь высоким коэффициентом диффузии в германии, при вплавлении диффундирует из германия в навеску. В результате из слоя германия, примыкающего к навеске, удаляется акцепторная примесь меди, изменяется тип электропроводности этого слоя, т. е. происходит так называемая *конверсия полупроводника*.

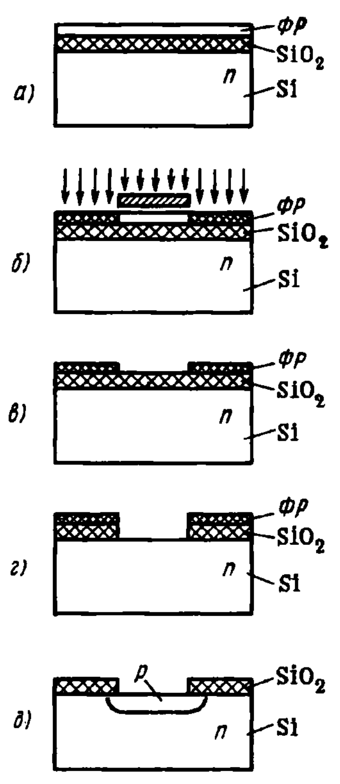


Рис. 2.5. Схема технологического процесса формирования планарного p-n-перехода:

а — окисленная пластинка полупроводника с нанесенным фоторезистом; *б* — экспонирование фоторезиста через маску или фотошаблон; *в* — пластинка после растворения незаполимеризованного фоторезиста; *г* — пластинка после вытравливания незащищенного слоя оксида; *д* — пластинка после снятия фоторезиста и проведения диффузии акцептора

При производстве полупроводниковых приборов широко используют *эпитаксиальное наращивание* — наращивание монокристаллических слоев полупроводника на поверхности монокристаллической подложки того же полупроводника, а иногда и другого по химическому составу полупроводника. При эпитаксиальном наращивании в зависимости от использованной примеси можно получить эпитаксиальный слой с тем же типом электропроводности, что и исходный полупроводник, но с другим удельным сопротивлением, а можно получить эпитаксиальный слой с другим типом электропроводности, т. е. создать *эпитаксиальный p-n-переход.*

Перспективным методом формирования p-n-переходов является метод *ионного внедрения* или *ионной имплантации*. Суть этого метода состоит в бомбардировке полупроводника ионами примеси с энергией в несколько десятков килоэлектрон-вольт. Необходимую энергию ионы получают при ускорении в электрическом поле ионно-лучевого ускорителя. Перспективность метода ионного внедрения заключается в возможности проводить управляемое легирование поверхностных и подповерхностных слоев полупроводника точно дозированными количествами почти любых химических элементов при относительно низкой температуре полупроводника.

По характеру распределения концентрации примеси различают резкие и плавные p-n-переходы. Переход, в котором толщина области изменения концентрации примеси значительно меньше толщины p-n-перехода, называют *резким p-n-переходом*. Резкий p-n-переход получается обычно при методе вплавления примеси. Переход, в котором толщина области изменения концентрации примеси сравнима или больше толщины p-n-перехода, называют *плавным p-n-переходом*. Плавный переход получают обычно при изготовлении методом диффузии примеси.

По соотношению концентраций основных носителей заряда или соответствующих примесей в p- и n-областях различают симметричные и несимметричные p-n-переходы. У *симметричных p-n-переходов* концентрации основных носителей заряда в прилегающих к переходу p и n-областях приблизительно равны ($p_{p0} \approx \approx n_{n0}$). Для *несимметричных p-n-переходов* справедливо неравенство $p_{p0} \gg n_{n0}$ (или $n_{n0} \gg p_{p0}$). В полупроводниковых приборах обычно существуют несимметричные p-n-переходы. Для обозначения несимметричных p-n-переходов пользуются следующими символами: p^+-n (или n^+-p).

§ 2.5. РАСПРЕДЕЛЕНИЕ НАПРЯЖЕННОСТИ ЭЛЕКТРИЧЕСКОГО ПОЛЯ И ПОТЕНЦИАЛА В ЭЛЕКТРОННО-ДЫРОЧНОМ ПЕРЕХОДЕ

При отсутствии токов распределение напряженности электрического поля и потенциала в p-n-переходе можно найти, решив *уравнение Пуассона*

$$\nabla^2\varphi = -\frac{q}{\varepsilon\varepsilon_0}(p - n + N_д - N_а). \qquad (2.9)$$

Однако даже для одномерной системы решение такого уравнения с учетом носителей заряда затруднено. Поэтому обычно влиянием носителей заряда на распределение потенциала пренебрегают, считая, что в p-n-переходе они отсутствуют, в то время как прилегающие к p-n-переходу области полупроводника остаются электрически нейтральными. Такие допущения соответству-

ют обратному и малому прямому напряжениям иа p-n-переходе при малых токах через него.

Тогда для плоского одномерного p-n-перехода распределение потенциала определяется уравнением

$$\frac{d^2\varphi}{dx^2} = -\frac{q}{\varepsilon\varepsilon_0} N(x), \qquad (2.10)$$

где $N(x) = N_д - N_а$ — разность концентраций доноров и акцепторов.

Необходимые для решения уравнения граничные условия можно записать с учетом следующих соображений. На одной границе p-n-перехода, где $x = -\delta_p$ (см. рис. 2.1), потенциал p-области примем за нуль, так как одно значение потенциала может быть взято произвольно. В связи с тем что напряжеиность электрического поля при отсутствии заряженных поверхностей изменяется непрерывно, на границе p-n-перехода она должна быть такой же, как и вне его. Практически в нейтральных областях полупроводника напряженность поля всегда значительно меньше, чем в p-n-переходе, и может быть принята равной нулю. На другой границе p-n-перехода при $x = \delta_n$ значение потенциала представляет собой алгебраическую сумму контактной разности потенциалов и внешнего напряжения u. Так как прямое напряжение вычитается из контактной разности потенциалов, а обратное — прибавляется, то $\varphi(\delta_n) = \varphi_{кон} - u$. Таким образом, граничные условия для уравнения (2.10) примут вид

$$\begin{aligned} &\text{при } x = -\delta_p \;\; \varphi = 0, \;\; \frac{d\varphi}{dx} = 0; \\ &\text{при } x = \delta_n \;\; \varphi = \varphi_{кон} - u, \;\; \frac{d\varphi}{dx} = 0. \end{aligned} \qquad (2.11)$$

Количество граничных условий оказалось больше, чем порядок уравнения (4 и 2). Это связано с тем, что координаты перехода $-\delta_p$ и δ_n неизвестны и для их нахождения потребуются дополнительные условия.

Распределение напряженности

Перепишем уравнение (2.10) в следующем виде:

$$\frac{d}{dx}\left(\frac{d\varphi}{dx}\right) = -\frac{q}{\varepsilon\varepsilon_0} N(x), \quad \text{тогда} \quad d\left(\frac{d\varphi}{dx}\right) = -\frac{q}{\varepsilon\varepsilon_0} N(x)dx.$$

Для упрощения записи в дальнейшем будем пользоваться определенными интегралами с переменными пределами. Это позволяет не применять произвольные постоянные, если пределы интегрирования согласованы. Тогда

$$\int_0^{d\varphi/dx} d\left(\frac{d\varphi}{dx}\right) = -\frac{q}{\varepsilon\varepsilon_0} \int_{-\delta_p}^{x} N(x)dx,$$

или

$$E=-\frac{d\varphi}{dx}=\frac{q}{\varepsilon\varepsilon_0}\int\limits_{-\delta_p}^{x} N(x)\,dx. \qquad (2.12)$$

Соотношение (2.12) дает возможность найти распределение напряженности электрического поля или градиента потенциала в p-n-переходе с любым характером изменения в нем разности концентраций доноров и акцепторов. Если для этого уравнения учесть граничные условия (2.11), то

$$\int\limits_{-\delta_p}^{\delta_n} N(x)\,dx=0. \qquad (2.13)$$

Уравнение (2.13) является *условием электрической нейтральности p-n-перехода*, т. е. свидетельствует о том, что суммарный неподвижный заряд ионизированных примесей с одной стороны металлургического контакта равен суммарному неподвижному заряду ионизированных примесей с другой стороны металлургического контакта.

Распределение потенциала

Для расчета распределения потенциала надо проинтегрировать градиент потенциала, т. е.

$$\varphi=\int\limits_{-\delta_p}^{x}\frac{d\varphi}{dx}\,dx. \qquad (2.14)$$

После подстановки (2.12) в (2.14) имеем

$$\varphi=-\frac{q}{\varepsilon\varepsilon_0}\int\limits_{-\delta_p}^{x}\left[\int\limits_{-\delta_p}^{x} N(x)\,dx\right]dx.$$

Правую часть полученного выражения проинтегрируем по частям:

$$\int y\,dx=yx-\int x\,dy, \text{ где } y=\int\limits_{-\delta_p}^{x} N(x)\,dx \text{ и}$$

$$dy=d\left[\int\limits_{-\delta_p}^{x} N(x)\,dx\right]=N(x)\,dx|_{-\delta_p}^{x}.$$

Тогда

$$\varphi=-\frac{q}{\varepsilon\varepsilon_0}\left[x\int\limits_{-\delta_p}^{x} N(x)\,dx-\int\limits_{-\delta_p}^{x} xN(x)\,dx\right]. \qquad (2.15)$$

Если для этого уравнения, показывающего распределение потенциала в p-n-переходе, учесть граничные условия (2.11), то полное падение потенциала на p-n-переходе

$$\varphi_{\text{кон}}-u=\frac{q}{\varepsilon\varepsilon_0}\int\limits_{-\delta_p}^{\delta_n} xN(x)\,dx. \qquad (2.16)$$

Уравнение электрической нейтральности p-n-перехода (2.13) и уравнение для полного падения напряжения на p-n-переходе (2.16) представляют собой систему двух уравнений с двумя неизвестными — пределами интегрирования. Решение этой системы уравнений дает положение границ p-n-перехода (δ_p и δ_n), которые до сих пор оставались неопределенными.

Методы решения полученных уравнений

Численное решение. Численные методы решения уравнений для распределения напряжениости электрического поля и потенциала в p-n-переходе целесообразно применять при сложном распределении концентрации примеси. Для такого решения используют численный метод интегрирования. Решение начинают при произвольно взятой координате $x=-\delta_p$. Выполняют численное определение градиента потенциала $d\varphi/dx$. Этот интеграл сначала возрастает по абсолютному значению, а затем падает, проходит через нуль и изменяет знак. Значение координаты, при которой выполняется условие электрической нейтральности, определяет вторую границу p-n-перехода.

Однако при этом остается неизвестным значение напряжения. Для его определения проводят интегрирование градиента потенциала $d\varphi/dx$. Значение этого интеграла при $x=\delta_n$ и дает значение напряжения на p-n-переходе $\varphi_{кон}-u$. Таким образом, напряжение на p-n-переходе при численном методе решения получается не аргументом, а функцией. Если же требуется выполнить расчет для заданного напряжения, то производят корректировку координаты $-\delta_p$.

В ходе такого решения распределения напряженности электрического поля и потенциала в p-n-переходе получаются как промежуточные значения интегралов.

Аналитическое решение. Аналитическое решение возможно в тех случаях, когда распределение концентрации примесей является интегрируемой аналитически функцией и полученная система уравнений имеет аналитическое решение. В частности, аналитическое решение возможно для резкого p-n-перехода, для плавного p-n-перехода с линейным распределением концентрации примесей и других плавных p-n-переходов.

§ 2.6. АНАЛИТИЧЕСКИЙ РАСЧЕТ РЕЗКОГО ЭЛЕКТРОННО-ДЫРОЧНОГО ПЕРЕХОДА

В соответствии с определением резкого p-n-перехода (см. § 2.4) можно считать, что концентрация примесей изменяется скачком на металлургическом контакте, а в остальных частях перехода остается постоянной, т. е.

$$N(x)=-N_{ap} \text{ при } -\delta_p<x<0$$

и

$$N(x)=N_{дn} \text{ при } \delta_n>x>0. \qquad (2.17)$$

Распределение напряженности

При $x<0$ в соответствии с уравнением (2.12) и с условиями (2.17) градиент потенциала

$$\left(\frac{d\varphi}{dx}\right)_p=\frac{q}{\varepsilon\varepsilon_0}N_{ap}(x+\delta_p). \qquad (2.18)$$

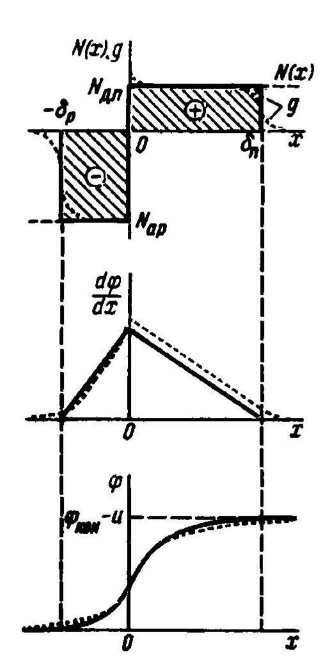


Рис. 2.6. Распределение концентрации примесей $N(x)$, плотности объемного заряда g, градиента потенциала $d\varphi/dx$ и потенциала φ в резком p-n-переходе без учета носителей заряда (——) и с учетом носителей заряда (...)

При $x>0$ следует учесть, что подынтегральная функция в (2.12) имеет разрыв. Поэтому целесообразно записать

$$\left(\frac{d\varphi}{dx}\right)_n=-\frac{q}{\varepsilon\varepsilon_0}\left[\int_{-\delta_p}^{0}N(x)dx+\int_{0}^{x}N(x)dx\right]=$$

$$=\frac{q}{\varepsilon\varepsilon_0}(N_{ap}\delta_p-N_{дл}x). \qquad (2.19)$$

Таким образом, для резкого p-n-перехода получается линейная зависимость градиента потенциала или напряженности электрического поля в p-n-переходе, наибольшее значение градиента получается на металлургическом контакте (рис. 2.6).

Распределение потенциала

Для расчета распределения потенциала надо проинтегрировать выражения для его градиента, т. е. (2.18) и (2.19). Тогда

$$\varphi_p=\frac{q}{\varepsilon\varepsilon_0}N_{ap}\int_{-\delta_p}^{x}(x+\delta_p)dx=$$

$$=\frac{q}{2\varepsilon\varepsilon_0}N_{ap}(x+\delta_p)^2 \qquad (2.20)$$

и

$$\varphi_n=\int_{-\delta_p}^{0}\left(\frac{d\varphi}{dx}\right)_p dx+\int_{0}^{x}\left(\frac{d\varphi}{dx}\right)_n dx=\frac{q}{2\varepsilon\varepsilon_0}N_{ap}\delta_p^2+\frac{q}{\varepsilon\varepsilon_0}N_{ap}\delta_p x-$$

$$-\frac{q}{2\varepsilon\varepsilon_0}N_{дл}x^2. \qquad (2.21)$$

Таким образом, потенциал в резком p-n-переходе изменяется с координатой по квадратичной параболе (рис. 2.6).

Отношение δ_p/δ_n

Полученные выражения (2.18)—(2.21) содержат пока не известные значения границ *p-n*-перехода. Воспользовавшись условием электрической нейтральности (2.13), для резкого *p-n*-перехода получим

$$\int_{-\delta_p}^{0}(-N_{ap})dx+\int_{0}^{\delta_n}N_{дn}dx=0;$$

$$\delta_p/\delta_n=N_{дn}/N_{ap}, \qquad (2.22)$$

т. е. металлургический контакт делит резкий *p-n*-переход на части, толщины которых обратно пропорциональны концентрациям примесей в соответствующих областях. Следовательно, при несимметричном легировании областей (что чаще всего и бывает) *p-n*-переход расположен преимущественно в области с меньшей концентрацией примесей.

Определение границ и толщины *p-n*-перехода

Для определения границ резкого *p-n*-перехода целесообразно ввести полную толщину перехода:

$$\delta=\delta_p+\delta_n.$$

Полную толщину *p-n*-перехода можно выразить через толщину одной его части, учитывая (2.22):

$$\delta=\delta_p+\delta_p\frac{N_{ap}}{N_{дn}}=\delta_p\frac{N_{ap}+N_{дn}}{N_{дn}};\quad \delta=\delta_n\frac{N_{дn}}{N_{ap}}+\delta_n=$$

$$=\delta_n\frac{N_{ap}+N_{дn}}{N_{ap}}.$$

Отсюда

$$\delta_p=\delta\frac{N_{дn}}{N_{ap}+N_{дn}}\ \text{ и }\ \delta_n=\delta\frac{N_{ap}}{N_{ap}+N_{дn}}. \qquad (2.23)$$

Найдя полную толщину *p-n*-перехода, можно с помощью соотношений (2.23) определить координаты его границ. Для получения выражения полной толщины *p-n*-перехода запишем полное падение потенциала после интегрирования (2.16) с учетом условий (2.17):

$$\varphi_{кон}-u=\frac{q}{2\varepsilon\varepsilon_0}(N_{ap}\delta_p^2+N_{дn}\delta_n^2). \qquad (2.24)$$

Подставив в (2.24) соотношения (2.23), получим

$$\varphi_{кон}-u=\frac{qN_{ap}N_{дn}\delta^2}{2\varepsilon\varepsilon_0(N_{ap}+N_{дn})}. \qquad (2.25)$$

Отсюда полная толщина резкого p-n-перехода

$$\delta = \sqrt{\frac{2\varepsilon\varepsilon_0}{q} \frac{N_{ap} + N_{дn}}{N_{ap} N_{дn}} (\varphi_{кон} - u)}. \qquad (2.26)$$

Реальные резкие p-n-переходы (например, сплавные) обычно бывают несимметричными, т. е.

$$N_{ap} \gg N_{дn} \text{ или } N_{дn} \gg N_{ap}.$$

Поэтому толщина резкого несимметричного p-n-перехода

$$\delta = \sqrt{\frac{2\varepsilon\varepsilon_0}{q} \frac{\varphi_{кон} - u}{N}}, \qquad (2.27)$$

где N — концентрация примесей в слаболегированной области.

Хотя в подкоренном выражении и стоит разность $\varphi_{кон} - u$, отрицательные значения этой разности не имеют физического смысла. При обратных напряжениях $u < 0$ и подкоренное выражение всегда положительно. При прямых напряжениях $u > 0$, но оно не может превзойти $\varphi_{кон}$, так как при $u \to \varphi_{кон}$ p-n-переход исчезает.

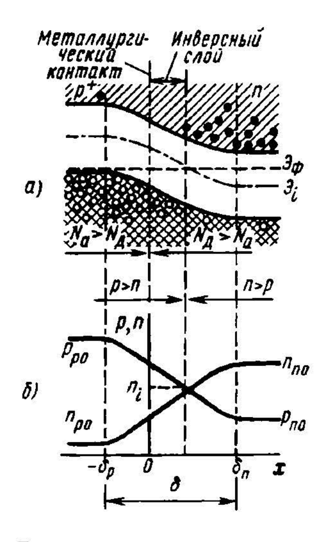


Рис. 2.7. Энергетическая диаграмма несимметричного резкого p-n-перехода в состоянии равновесия с инверсным слоем в слаболегированной области перехода (*а*) и распределение носителей заряда в таком переходе (*б*)

Влияние носителей заряда

Подвижные носители заряда могут оказывать влияние на распределение градиента потенциала и потенциала в резком p-n-переходе вблизи его границ и вблизи металлургического контакта.

Вблизи границ p-n-перехода с учетом предположений, принятых для расчетов, плотность объемного заряда нескомпенсированных примесей изменяется скачком от нуля за пределами p-n-перехода до значения, отличающегося только множителем q от $N(x)$, т. е. в пределах p-n-перехода график распределения плотности объемного заряда отличается от графика распределения концентрации примесей только масштабом (рис. 2.6). В действительности же плотность объемного заряда не может изменяться скачком, так как это соответствовало бы бесконечному градиенту концентрации носителей заряда и наличию бесконечно больших диффузион-

ных токов. Поэтому границы p-n-перехода должны быть размытыми, а в распределениях градиента потенциала и потенциала должны быть «хвосты» (рис. 2.6).

Вблизи металлургического контакта несимметричного резкого p-n-перехода концентрация носителей заряда также может существенно отличаться от собственной, что иллюстрируется энергетической диаграммой p-n-перехода в состоянии равновесия (рис. 2.7). При несимметрии p-n-перехода металлургический контакт оказывается ближе к сильнолегированной области, а падение потенциала там меньше. Уровень Ферми пересекает середину запрещенной зоны не на металлургическом контакте, а в слаболегированной части p-n-перехода. Тип преобладающих носителей заряда определяется взаимным положением уровня Ферми и середины запрещенной зоны: где уровень Ферми выше середины запрещенной зоны, преобладают электроны, где ниже — дырки. Следовательно, граница смены типа носителей заряда не совпадает с металлургическим контактом, т. е. в некотором слое слаболегированной части p-n-перехода тип преобладающих (основных) носителей заряда не соответствует типу примесей. Такой слой, как указывалось (см. § 1.13), называют инверсным.

§ 2.7. АНАЛИТИЧЕСКИЙ РАСЧЕТ ПЛАВНОГО ЭЛЕКТРОННО-ДЫРОЧНОГО ПЕРЕХОДА С ЛИНЕЙНЫМ РАСПРЕДЕЛЕНИЕМ КОНЦЕНТРАЦИИ ПРИМЕСЕЙ

Линейный закон распределения примесей

$$N(x)=ax, \qquad (2.28)$$

здесь a — градиент концентрации примесей, который можно считать постоянным при очень малой толщине p-n-перехода по сравнению с толщиной области, где концентрация примесей переменна.

Распределение напряженности

Решение уравнения (2.12) с учетом (2.28) дает зависимость градиента потенциала в плавном p-n-переходе с линейным распределением концентрации примесей:

$$\frac{d\varphi}{dx}=-\frac{q}{2\varepsilon\varepsilon_0}ax^2\Big|_{-\delta_p}^{x}=-\frac{qa}{2\varepsilon\varepsilon_0}(x^2-\delta_p^2). \qquad (2.29)$$

В данном случае градиент потенциала изменяется с изменением координаты по квадратичной параболе (рис. 2.8).

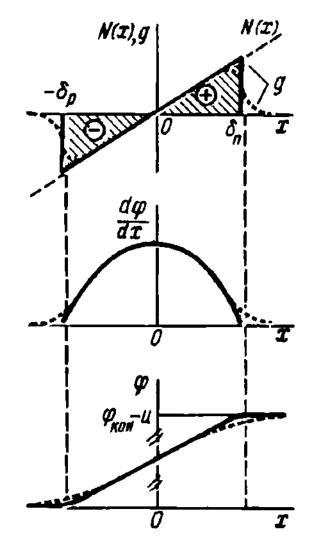


Рис. 2.8. Распределение концентрации примесей $N(x)$, плотности объемного заряда g, градиента потенциала $d\varphi/dx$ и потенциала φ в плавном p-n-переходе без учета носителей заряда (—) и с учетом носителей заряда (...)

Распределение потенциала

Для определения потенциала надо проинтегрировать градиеит потенциала (2.29):

$$\varphi = -\frac{qa}{2\varepsilon\varepsilon_0}\int_{-\delta_p}^{x}(x^2-\delta_p^2)dx = -\frac{qa}{\varepsilon\varepsilon_0}\left(\frac{x^3}{6}-\frac{\delta_p^2 x}{2}-\frac{\delta_p^3}{3}\right). \quad (2.30)$$

Таким образом, потенциал в плавном p-n-переходе с линейным распределением примесей изменяется с изменением координаты по кубической параболе (рис. 2.8).

Определение границ и толщины p-n-перехода

Решение уравнения электрической нейтральности (2.13) в данном случае приводит к соотношению

$$\delta_p = \delta_n = \delta/2, \quad (2.31)$$

т. е. рассматриваемый переход является симметричным.

Из уравнения (2.16) с учетом (2.28) можно получить полное падение потенциала на p-n-переходе:

$$\varphi_{\text{кон}} - u = \varphi(\delta_n) - \varphi(\delta_p) = \frac{2qa}{3\varepsilon\varepsilon_0}\delta_p^3 = \frac{qa}{12\varepsilon\varepsilon_0}\delta^3. \quad (2.32)$$

Отсюда толщина плавного p-n-перехода с линейным распределением примесей

$$\delta = \sqrt[3]{\frac{12\varepsilon\varepsilon_0(\varphi_{\text{кон}} - u)}{qa}}. \quad (2.33)$$

Влияние носителей заряда

В плавном p-n-переходе с линейным распределением концентрации примесей из-за симметрии перехода инверсный слой не образуется, если нет существенной несимметрии плотности энергетических состояний в разрешенных энергетических зонах, т. е. если нет резкого отличня эффективных масс носителей заряда разных знаков.

У границ плавного p-n-перехода влияние носителей заряда аналогично влиянню носителей у резкого p-n-перехода.

§ 2.8. БАРЬЕРНАЯ ЕМКОСТЬ ЭЛЕКТРОННО-ДЫРОЧНОГО ПЕРЕХОДА

Барьерная емкость как проявление токов смещения

Барьерная емкость p-n-перехода проявляется при приложении к p-n-переходу изменяющегося во времени напряжения. При этом через p-n-переход проходит ток. Та доля тока, которая не связана с движением носителей заряда через p-n-переход, и определяет барьерную емкость.

Поэтому барьерная емкость должна быть связана с прохождением токов смещения. Для одномерного плоского p-n-перехода ток смещения одинаков во всех его сечениях:

$$i = SJ_{\text{см}} = S\frac{d}{dt}(\varepsilon\varepsilon_0 E),$$

где S — площадь p-n-перехода.

Значение тока смещения можно связать с изменением объемного заряда. Для этого выделим мысленно в p-n-переходе объем в виде цилиндра (или призмы), образующие которого параллельны оси x — направлению электрического поля (рис. 2.9). Пусть одно основание цилиндра лежит за пределами p-n-перехода, а другое — внутри его. Тогда, согласно теореме Остроградского—Гаусса, можно определить поток вектора электрической индукции через поверхность, ограничивающую выделенный объем. Этот поток проходит только через одно основание цилиндра, так как боковые его поверхности параллельны электрическому полю, а второе основание лежит в области, где поле отсутствует. Следовательно,

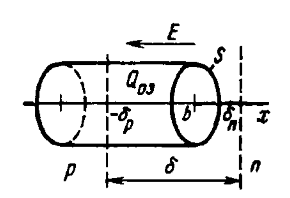


Рис. 2.9. К выводу выражения для барьерной емкости p-n-перехода

$$\varepsilon\varepsilon_0 ES = Q,$$

где Q — заряд ионизированных примесей.

Ток смещения можно записать теперь таким образом:

$$i = \frac{dQ}{dt} = \frac{dQ}{dU}\frac{dU}{dt}.$$

Сравнив последнее выражение с обычным выражением для тока через емкость, т. е. с

$$i_C = S\frac{dU}{dt},$$

получаем, что в качестве барьерной емкости следует взять

$$C_{\text{бар}} = \left|\frac{dQ}{dU}\right|. \tag{2.34}$$

Абсолютное значение этого отношения взято потому, что объемный заряд в p-n-переходе может быть положительным и отрицательным, а правило знаков для напряжения выбрано произвольно.

Таким образом, барьерная емкость связана с током смещения (как и обычная емкость).

Общее соотношение для барьерной емкости электронно-дырочного перехода

Исходя из определения барьерной емкости (2.34) можно получить общую формулу для барьерной емкости плоского p-n-перехода. Объемный заряд ионизированных примесей в цилиндре, выделенном в p-n-переходе,

$$Q = qS\int_{-\delta_p}^{b} N(x)dx.$$

Дифференциал этого объемного заряда можно определить путем дифференцирования по единственной переменной — нижнему пределу интегрирования:

$$dQ = -qSN(-\delta_p)d(-\delta_p). \tag{2.35}$$

Приращение напряжения или дифференциал падения напряжения на p-n-переходе можно найти, продифференцировав выражение (2.16). Но при этом следует учесть, что переменными являются оба предела интегрирования в выражении (2.16). Поэтому целесообразно разбить интеграл на два, тогда каждый из них будет иметь по одному переменному пределу:

$$\varphi_{кон} - U = \frac{q}{\varepsilon\varepsilon_0}\int_{-\delta_p}^{\delta_n} xN(x)dx = \frac{q}{\varepsilon\varepsilon_0}\left[\int_{-\delta_p}^{0} xN(x)dx + \int_{0}^{\delta_n} xN(x)dx\right]$$

и

$$dU = \frac{q}{\varepsilon\varepsilon_0}[\delta_p N(-\delta_p)d(-\delta_p) + \delta_n N(\delta_n)d\delta_n]. \tag{2.36}$$

Теперь, взяв отношение дифференциалов (2.35) и (2.36), получим

$$C_{бар} = \left|\frac{qSN(-\delta_p)d\delta_p}{\frac{q}{\varepsilon\varepsilon_0}[\delta_p N(-\delta_p)d(-\delta_p) + \delta_n N(\delta_n)d\delta_n]}\right|.$$

Для преобразования последнего выражения продифференцируем условие электрической нейтральности p-n-перехода (2.13):

$$N(\delta_n)d\delta_n - N(-\delta_p)d(-\delta_p) = 0,$$

т. е.

$$N(\delta_n)d\delta_n = N(-\delta_p)d(-\delta_p).$$

Учтем, что $\delta = \delta_p + \delta_n$, тогда

$$C_{бар} = \varepsilon\varepsilon_0 S/\delta. \qquad (2.37)$$

Таким образом, барьерная емкость плоского одномерного p-n-перехода может быть рассчитана по формуле плоского конденсатора. Такой результат не является очевидным, так как распределение зарядов в плоском конденсаторе и в электронно-дырочном переходе не одинаково. Причина совпадения формул — в характере изменения заряда p-n-перехода: при изменении напряжения на p-n-переходе заряд изменяется потому, что сдвигаются границы p-n-перехода. Заряды, обусловливающие барьерную емкость, сосредоточены в двух тонких слоях, расположенных на расстоянии δ один от другого (рис. 2.10), что очень напоминает поверхностные заряды на металлических обкладках конденсатора.

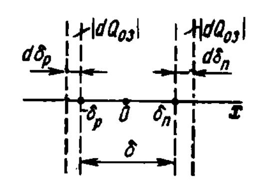


Рис. 2.10. Изменение толщины p-n-перехода и объемного заряда при изменении напряжения на p-n-переходе

Частные соотношения для барьерной емкости различных электронно-дырочных переходов

С помощью выражения (2.37) можно определить барьерную емкость на основе результатов расчета толщины p-n-перехода. Поэтому для резкого p-n-перехода с учетом (2.26)

$$C_{бар} = S\sqrt{\frac{q\varepsilon\varepsilon_0 N_{ap} N_{дn}}{2(N_{ap} + N_{дn})(\varphi_{кон} - U)}}. \qquad (2.38)$$

Для резкого несимметричного p-n-перехода с учетом (2.27)

$$C_{бар} = S\sqrt{\frac{q\varepsilon\varepsilon_0 N}{2(\varphi_{кон} - U)}}, \qquad (2.39)$$

где N — концентрация примесей в слаболегированной области.

Для плавного p-n-перехода с линейным распределением концентрации примесей с учетом (2.33)

$$C_{бар} = S\sqrt[3]{\frac{(\varepsilon\varepsilon_0)^2 qa}{12(\varphi_{кон} - U)}}. \qquad (2.40)$$

Как видно из полученных результатов, при разных распределениях примесей имеют место разные вольт-фарадные характеристики p-n-переходов. Это дает возможность оценивать характер распределения примесей в различных p-n-переходах. Часто применяют также графический метод. Для резкого p-n-перехода вольт-фарадная характеристика оказывается прямой в координатах $1/C^2_{бар}$ от U, а для плавного перехода с линейным распределением примесей — в координатах $1/C^3_{бар}$ от U (рис. 2.11).

Если экспериментальные точки ложатся на прямые в указанных системах координат, то это служит подтверждением (но не доказательством) принятого при построении характера распределения примесей. Однако вольт-фарадные характеристики неоднозначно связаны с распределением примесей в *p-n*-переходе, т. е. одинаковым вольт-фарадным характеристикам могут соответствовать разные распределения примесей. Например, распределения, показанные на рис. 2.12, дают прямые вольт-фарадные характеристики в координатах $1/C_{бар}^3$ от U. Поэтому при анализе вольт-фарадных характеристик необходимо принимать во внимание дополнительные сведения о технологии и т. п.

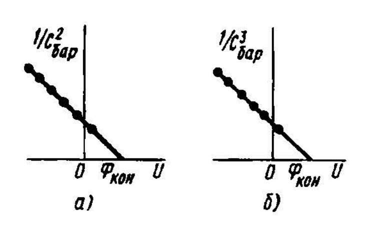


Рис. 2.11. Зависимость барьерной емкости резкого (*а*) и плавного (*б*) *p-n*-переходов от постоянного смещения на переходе

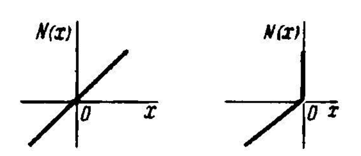


Рис. 2.12. Различные распределения примесей, обеспечивающие одинаковые вольт-фарадные характеристики *p-n*-переходов

Кроме того, вольт-фарадные характеристики дают возможность определить значение контактной разности потенциалов на *p-n*-переходе (или высоту потенциального барьера). При экстраполяции вольт-фарадной характеристики, построенной в координатах $1/C_{бар}^2=(U)$ или $1/C_{бар}^3=(U)$, отрезок, отсекаемый ею на оси положительных напряжений, соответствует значению контактной разности потенциалов (см. рис. 2.11). Таким образом, рассмотренный метод определения контактной разности потенциалов основан на том, что при постоянном напряжении смещения, стремящемся к контактной разности потенциалов, барьерная емкость стремится к бесконечности. Необходимость экстраполяции вольт-фарадной характеристики связана с малой добротностью, т. е. с большими прямыми токами, проходящими через *p-n*-переход при больших прямых напряжениях, и, следовательно, с практической невозможностью точного определения барьерной емкости перехода. Для удобства экстраполяции необходимо выбрать координаты, в которых вольт-фарадная характеристика исследуемого перехода соответствует прямой линии.

§ 2.9. ОМИЧЕСКИЙ ПЕРЕХОД НА КОНТАКТЕ ПОЛУПРОВОДНИКОВ С ОДНИМ ТИПОМ ЭЛЕКТРОПРОВОДНОСТИ

При контакте двух полупроводников с электропроводностью одного типа, но с различными значениями удельной электрической проводимости происходят процессы, аналогичные процессам

в p-n-переходе, т. е. носители заряда из области с большей концентрацией диффундируют в область с меньшей концентрацией. В сильнолегированной области нарушается компенсация заряда ионизированных примесей, а в слаболегированной — создается избыток основных носителей заряда (рис. 2.13, *в*). Таким образом, на контакте двух полупроводников с электропроводностью одного типа, но с различными значениями удельной электрической проводимости также образуются область пространственного заряда, диффузионное электрическое поле и контактная разность потенциалов. Но в отличие от p-n-перехода в данном случае в слаболегированной области объемный заряд образуется в результате избыточной концентрации основных носителей заряда.

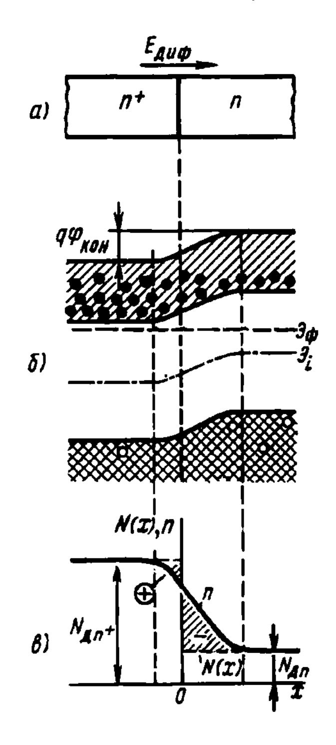


Рис. 2.13. Омический переход между полупроводниками с одним типом электропроводности:

а — структура перехода; *б* — энергетическая диаграмма; *в* — зависимости концентрации примесей и концентрации носителей заряда от координаты

Переходный слой в полупроводнике между двумя областями с электропроводностью n-типа, обладающими различными значениями удельной электрической проводимости, в котором существует диффузионное электрическое поле, называют ***электронно-электронным переходом*** (n-n^+-переходом).

Аналогичные процессы происходят вблизи границы раздела двух областей полупроводника с электропроводностью p-типа. Переходный слой в полупроводнике между двумя областями с электропроводностью p-типа, обладающими различными значениями удельной электрической проводимости, в котором существует диффузионное электрическое поле, называют ***дырочно-дырочным переходом*** (p-p^+-переходом). Знак «+» (плюс) в этих терминах условно означает область с более высокой удельной электрической проводимостью, т. е. с большей концентрацией соответствующей примеси.

Контактная разность потенциалов на таких переходах определяется соотношениями, аналогичными (2.3):

для n-n^+-перехода

$$\varphi_{кон} = \frac{kT}{q} \ln \frac{n_{n^+0}}{n_{n0}}; \qquad (2.41)$$

для p-p^+-перехода

$$\varphi_{\text{кон}} = \frac{kT}{q} \ln \frac{p_{p^+0}}{p_{p0}}. \qquad (2.42)$$

Распределение носителей заряда в n-n^+-переходе (рис. 2.13, *в*) показывает, что в отличие от p-n-перехода в данном случае отсутствует обедненный слой — слой с меньшей концентрацией основных носителей заряда по сравнению с концентрацией носителей заряда в слаболегированной области. При приложении внешнего напряжения к структуре с n-n^+-переходом практически все напряжение падает на слаболегированной (высокоомной) области. Поэтому высота потенциального барьера на n-n^+-переходе не зависит от полярности приложениого напряжения и от его значения. Таким образом, n-n^+-переход (и p-p^+-переход) имеет малое сопротивление по сравнению с сопротивлением слаболегированной области и не обладает выпрямляющими свойствами. Переход, электрическое сопротивление которого не зависит от направления тока в заданном диапазоне значений токов, называют *омическим переходом.*

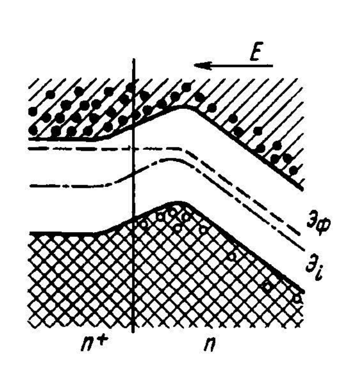


Рис. 2.14. Накопление неосновиых иосителей заряда (дырок) вблизи омического перехода между полупроводинкамн с электропроводностью n-типа при иаличии виешиего электрического поля

Другой особенностью n-n^+-перехода (и p-p^+-перехода) является отсутствие инжекции неосновных носителей заряда в слаболегированную область. Действительно, если внешнее напряжеиие приложено положительным потенциалом к n-области n-n^+-перехода, что аналогично прямому включению p-n-перехода, то из n^+-области в n-область вводятся электроны, которые являются основными носителями заряда. При противоположной полярности дырочный ток из n^+-области в n-область аналогичеи обратному току через p-n-переход. Однако из-за ничтожно малой концентрации неосиовных носителей заряда в сильнолегированной n^+-области [см. (1.19)] поток дырок в высокоомную n-область также оказывается ничтожно малым.

Омические и неинжектирующие переходы широко используют в полупроводииковых приборах наряду с выпрямляющими и иижектирующими.

Однако из-за существования иа n-n^+-переходе (и на p-p^+-переходе) потенциального барьера для иеосновных носителей заряда, двигающихся из слаболегированной области к переходу, может происходить накопление этих неосновиых носителей вблизи перехода при определенной полярности внешиего напряжеиия

(рис. 2.14). При приложении внешнего напряжения оно падает в основном в объеме слаболегированной области, где возникает наклон энергетических уровней и зон, приводящий к образованию потенциальной ямы для иеосновных носителей заряда. Эффект накопления неосновных носителей заряда и последующего их рассасывания — эффект инерционный. Поэтому он может ухудшать быстродействие полупроводниковых приборов.

§ 2.10. ВЫПРЯМЛЯЮЩИЕ И ОМИЧЕСКИЕ ПЕРЕХОДЫ НА КОНТАКТЕ МЕТАЛЛА С ПОЛУПРОВОДНИКОМ

При идеальном контакте металла с полупроводником (т. е. при отсутствии каких-либо промежуточных слоев, отличающихся химическим составом, и без учета поверхностных состояний на границе раздела) происходит диффузия электронов преимущественио из материала с меньшей работой выхода электронов в материал с большей работой выхода. Под *работой выхода электронов* будем понимать энергию, необходимую для перевода электрона с уровня Ферми иа потолок верхней свободной зоны (без удаления электрона в вакуум на бесконечное расстояние от поверхиости полупроводника).

В результате диффузии электронов и перераспределения зарядов нарушается электрическая нейтральность прилегающих к границе раздела областей, возникает контактное электрическое поле и контактная разиость потенциалов

$$\varphi_{кон}=(A_м-A_п)/q, \tag{2.43}$$

где $A_м$ и $A_п$ — соответственно работа выхода электронов из металла и из полупроводника.

Переходиый слой, в котором существует контактное (или диффузиониое) электрическое поле и который образован в результате коитакта между металлом и полупроводником, называют *переходом Шотки,* по имени немецкого учеиого В. Шотки, который первым получил осиовиые математические соотиошения для электрических характеристик переходов.

Контактное электрическое поле на переходе Шотки сосредоточено практически только в полупроводнике, так как коицентрация носителей заряда в металле значительно больше концентрации носителей заряда в полупроводнике. Перераспределение электронов в металле происходит в очень тонком слое, сравнимом с межатомиым расстоянием.

В зависимости от типа электропроводности полупроводника и от соотношения работ выхода в полупроводнике может возникать обеднеиный, инверсиый или обогащенный слой (рис. 2.15). Если работа выхода в металле меньше работы выхода в полупровод-

нике ($A_м < A_п$), то электроны с большей вероятностью будут переходить из металла в полупроводник. Это приводит к образованию в полупроводнике обедненного слоя, если полупроводник *p*-типа (рис. 2.15, *а*), или даже инверсного слоя, если $A_м \ll A_п$ (рис. 2.15, *б*). Если полупроводник *n*-типа, то образуется обогащенный слой (рис. 2.15, *в*).

При противоположном соотношении работ выхода ($A_м > A_п$) в полупроводнике *n*-типа получается обедненный или инверсный слой, а в дырочном — обогащенный.

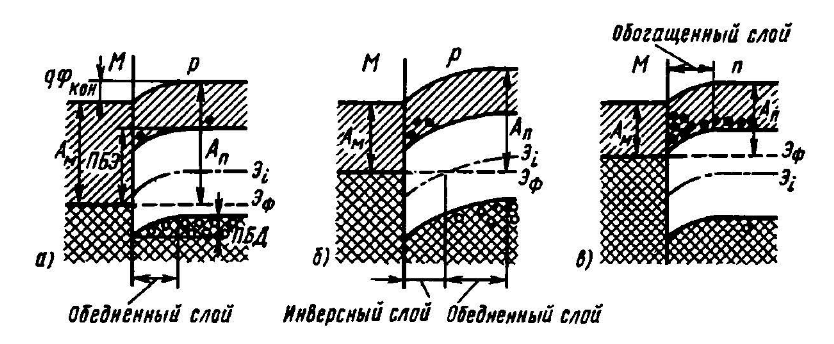


Рис. 2.15. Образование обедненного (*а*), инверсного (*б*) и обогащенного (*в*) слоев в полупроводнике вблизи металлургического контакта с металлом при работе выхода в металле меньше, чем в полупроводнике

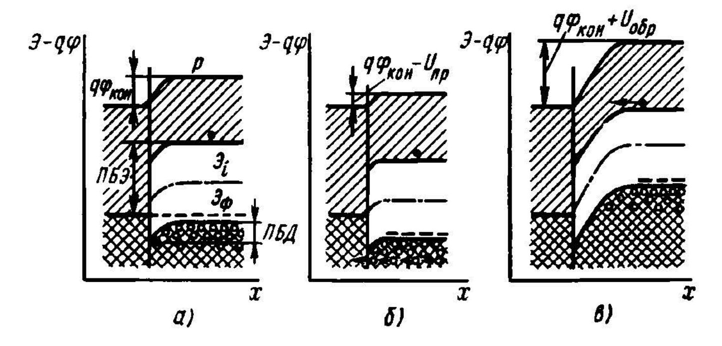


Рис. 2.16. Изменение высоты потенциального барьера на выпрямляющем неинжектирующем переходе между металлом и полупроводником при изменении внешнего напряжения:

а — внешнее напряжение отсутствует; *б* — внешнее напряжение прямое; *в* — внешнее напряжение обратное

В обедненных слоях пространственный заряд формируется в результате нарушения компенсации заряда ионизированных примесей основными носителями, а в обогащенных — из-за накопления основных носителей заряда. Обогащенный слой обусловливает малое сопротивление приконтактной области полупроводника по сравнению с сопротивлением объема полупроводника. По-

этому такой переход не обладает выпрямляющими свойствами. При наличии обедненного или инверсного слоя переход Шотки обладает выпрямляющими свойствами, так как внешнее напряжение, падая в основном на высокоомном переходе, будет изменять высоту его потенциального барьера, изменяя условия прохождения носителей заряда через переход.

Характерной особенностью выпрямляющего перехода Шотки в отличие от *p-n*-перехода является разная высота потенциальных барьеров для электронов и дырок. В результате через переход Шотки может не происходить инжекции неосновных носителей заряда в полупроводник. Рассмотрим рис. 2.16. При включении такого перехода в прямом направлении (рис. 2.16, *б*) высота потенциального барьера для дырок (ПБД) в приконтактной области полупроводника понижается, дырки будут переходить из полупроводника в металл. Чем больше прямое напряжение, тем больше вероятность такого перехода дырок. Однако при этом высота потенциального барьера для электронов (ПБЭ), которые могут двигаться из металла в полупроводник, остается еще относительно большой. Поэтому поток электронов из металла в полупроводник будет относительно малым, т. е. практически не будет инжекции неосновных носителей заряда в полупроводник.

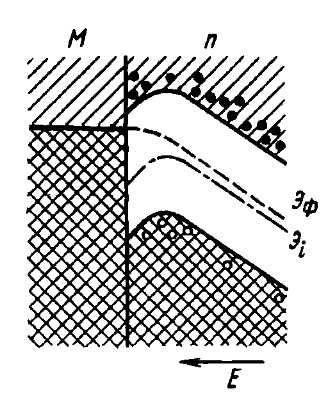


Рис. 2.17. Накопление неосновных носителей заряда (дырок) вблизи омического перехода между металлом и полупроводником при наличии внешнего электрического поля

При другой полярности внешнего напряжения (при обратном напряжении) потенциальный барьер для дырок повышается (рис. 2.16, *в*) и их движение через переход прекращается. Для неосновных носителей заряда (для электронов в данном примере) поле в переходе оказывается ускоряющим. Поэтому, проходя через переход, неосновные носители заряда образуют обратный ток, который будет мал из-за малой концентрации неосновных носителей в полупроводнике.

Если разница в работах выхода велика, то в приконтактной области полупроводника образуется инверсный слой (см. рис. 2.15, *б*). В этом случае при малых прямых напряжениях через такой переход будет происходить инжекция неосновных носителей заряда из инверсного слоя в прилегающий объем полупроводника. При больших прямых напряжениях инверсный слой может исчезнуть.

В омических переходах, образованных в результате контакта между металлом и полупроводником, может происходить накопление неосновных носителей заряда из-за образования потен-

циальной ямы для неосновных носителей в приконтакной области полупроводника (рис. 2.17).

Такое явление, как отмечалось в § 2.9, может влиять иа быстродействие полупроводниковых приборов. Для устранения этого явления необходимо устранить потенциальный барьер на контакте металл—полупроводник путем подбора контактирующих пар материалов с одинаковыми работами выхода. Одиако практически это неосуществимо из-за ограниченного набора материалов и необходимости подбирать металл заново для каждой концентрации примесей в полупроводнике и для каждой температуры.

Для исключения эффекта иакопления неосновных носителей заряда в полупроводнике около контакта можно провести дополиительное легирование приконтактной области полупроводника. Потеициальный барьер при этом остается, но его толщина будет очень малой из-за сильного легирования приконтактной области полупроводника. Малая толщина потеициального барьера обеспечит возможность туннелирования неосновных носителей заряда в металл из потенциальной ямы в полупроводнике.

§ 2.11. ГЕТЕРОПЕРЕХОДЫ

Гетеропереходом называют переходный слой с существующим там диффузионным электрическим полем между двумя различиыми по химическому составу полупроводниками.

При образовании гетероперехода из-за разных работ выхода электронов из разных полупроводников происходит перераспределение носителей заряда в приконтактной области и выравни-

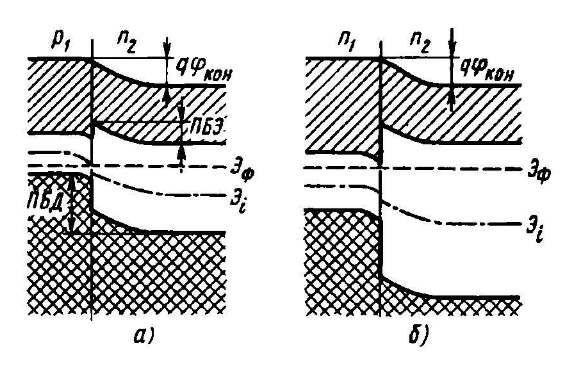


Рис. 2.18. Энергетические диаграммы гетеропереходов:

а — выпрямляющий гетеропереход между полупроводниками с электропроводностью *p*- и *n* типа с преимущественной инжекцией электронов в узкозонный полупроводник; *б* — выпрямляющий гетеропереход между полупроводниками с электропроводиостью *n*-тнпа без инжекции неосновиых носителей заряда

вание уровней Ферми в результате устаиовления термодииамического равновесия (рис. 2.18). Все остальные энергетические уровни и зоны должиы соответствеиио изогнуться, т. е. в гетеропереходе возникают диффузиоииое электрическое поле и контактиая разность потенциалов. При этом энергетический уровень потолка верхией свободной зоны должен быть непрерывным. Обычио энергетический уровеиь потолка верхней свободной зоны является энергетическим уровнем потолка зоны проводимости, так как свободиые эиергетические зоны перекрывают друг друга. Контактная разность потенциалов, возникающая иа гетеропереходе, определяется относительным смещением потолка верхней свободной зоны полупроводников, образующих гетеропереход.

Ширина эиергетических зон различных полупроводников различна. Поэтому на границе раздела двух полупроводников (на металлургическом контакте гетероперехода) получается обычно разрыв дна зоиы проводимости. Разрыв диа зоны проводимости определяется различием энергий сродства к электрону двух коитактирующих полупроводников (*энергия сродства к электрону* есть разница энергий потолка верхией свободной зоны и дна зоны проводимости).

Разрыв потолка валентной зоны зависит как от разницы энергий сродства, так и от различия ширины запрещенных зон контактирующих полупроводников.

В результате разрывов дна зоны проводимости и потолка валентиой зоны высота потеициальных барьеров для электронов и дырок в гетеропереходе оказывается различной. Это является особенностью гетеропереходов, обусловливающей специфические свойства гетеропереходов в отличие от *р-п*-переходов, которые формируются в монокристалле одного полупроводника.

Каждый из полупроводников, образующих гетеропереход, может иметь различный тип электропроводности. Поэтому для каждой пары полупроводников в принципе можно осуществить четыре разиовидности гетероструктур: p_1-n_2; n_1-n_2; n_1-p_2 и p_1-p_2.

Если вблизи границы раздела двух полупроводиико в, образующих гетеропереход, возникают обедненные основными носителями слои (слои с повышенным удельиым сопротивлением), то основная часть внешнего напряжения, приложенного к структуре с гетеропереходом, будет падать на обедненных слоях. Высота потенциального барьера для основных носителей заряда будет изменяться: уменьшаться при полярности внешнего напряжения, противоположной полярности контактной разности потенциалов, н увелнчиваться при совпадении полярностей внешнего напряжения и контактной разности потенциалов. Таким образом, гетеропереходы могут обладать эффектом выпрямления (рис. 2.18).

Из-за различия по высоте потенциальных барьеров для электронов и для дырок прямой ток через гетеропереход связаи в

основном с движением иосителей заряда только одиого знака. Поэтому гетеропереходы могут быть как инжектирующими неосновные иосители (рис. 2.18, *а*), так и неинжектирующими (рис. 2.18, *б*). Инжекция неосновных носителей заряда происходит всегда из широкозонного в узкозонный полупроводник. В гетеропереходах, образованных полупроводниками с одним типом электропроводности, выпрямление происходит без инжекции неосновных носителей заряда.

Обычио полупроводники различного химического состава отличаются друг от друга работой выхода электронов, шириной запрещениой зоны, шириной разрешенных зон и другими параметрами. Однако для формирования качественного гетероперехода необходимо совпадение типа, ориентации и периода кристаллических решеток контактирующих полупроводников, чтобы кристаллическая решетка одного полупроводника с минимальным количеством иарушеиий переходила в кристаллическую решетку другого полупроводника. В идеальиом гетеропереходе не должио быть механических напряжений, структурных и других дефектов, которые могут создать условия для интенсивной рекомбииации и генерации иосителей заряда — рекомбинационные ловушки. При наличии в гетеропереходе большого числа рекомбинационных ловушек механизм прохождения тока через такой реальиый гетеропереход может существеиио отличаться от механизма прохождения тока через идеальный гетеропереход. Такой гетеропереход может и не обладать эффектом выпрямления.

Кроме рассмотреиных специфических свойств гетеропереходов (выпрямление с высоким коэффициентом инжекции в узкозонный полупроводник, выпрямлеиие без инжекции неосновиых носителей на гетеропереходе из полупроводников с одним типом электропроводности) для полупроводниковых приборов интересными и полезными оказываются различия спектров поглощения и показателей преломления света образующих гетеропереход полупроводииков.

Наиболее широкое применеиие в полупроводниковых приборах имеют гетеропереходы между полупроводниками типа $A^{III}B^{V}$ и их твердыми растворами иа основе арсенидов, фосфидов и аитимонидов галлия и алюмииия. Благодаря близости ковалеитных радиусов галлия и алюминия изменеиие химического состава полупроводников в гетеропереходе происходит без измеиения периода кристаллической решетки. Гетеропереходы создают также на основе миогокомпонентных (четверных и более) твердых растворов, в которых при измеиении состава в широких пределах период решетки ие изменяется. Осиовным методом формирования гетероструктур является метод эпитаксиальиого наращивания полупроводниковых кристаллов.

Все рассмотренные ранее виды *электрических переходов* (p-n-переход, n-n^+-переход, p-p^+-переход, переход Шотки) можно считать частными случаями общего вида электрических переходов — гетероперехода.

§ 2.12. СВОЙСТВА И ПАРАМЕТРЫ ОМИЧЕСКИХ ПЕРЕХОДОВ

Требования к омическим переходам

Омические переходы имеют очень большое значение в полупроводниковых приборах и при проведении исследований полупроводников. Основное назначение омических переходов — электрическое соединение полупроводника с металлическими токопроводящими частями полупроводникового прибора. Омических переходов в полупроводниковых приборах больше, чем выпрямляющих. Случаи отказов и производственного брака полупроводниковых приборов из-за низкого качества омических переходов довольно часты. При разработке полупроводниковых приборов создание совершенных омических переходов нередко требует бо́льших усилий, чем создание выпрямляющих переходов.

Несмотря на это, теория омических переходов разработана слабее, чем теория электронно-дырочных переходов, а формирование омических переходов часто основано на эксперименте.

Омический переход оказывает меньшее отрицательное влияние на характеристики и параметры полупроводникового прибора при следующих условиях:

1) если отсутствуют инжекция неосновных носителей заряда через омический переход в прилегающую область полупроводника и накопление неосновных носителей в омическом переходе или вблизи него;

2) при минимально возможном падении напряжения на омическом переходе, т. е. при минимальном его сопротивлении;

3) если ВАХ омического перехода линейна, т. е. если омический переход действительно является омическим.

Параметры омических переходов

Чтобы можно было установить, насколько омический переход удовлетворяет предъявленным к нему требованиям, и чтобы можно было сравнивать разные омические переходы между собой, надо ввести количественные параметры, характеризующие эти переходы.

Скорость рекомбинации на омическом переходе. Скорость рекомбинации на омическом переходе показывает, насколько концентрация носителей заряда вблизи него может отклоняться от равновесной концентрации. Скорость рекомбинации на омическом переходе вводят аналогично скорости поверхностной рекомбинации [см. (1.36)] как отношение плотности потока носителей заряда через переход к избыточной концентрации этих носителей у перехода, т. е.

$$s_p = \Phi_p/(p_{гр} - p_0).$$

Размерность скорости рекомбинации на омическом переходе та же, что и размерность скорости движения, так как плотность потока носителей — это произведение концентрации на скорость ($\Phi_p = pv_p$). Очевидно, что чем больше скорость рекомбинации, тем меньше при данном потоке носителей заряда отклонение концентрации носителей от равновесной, тем выше качество омического перехода. При большой плотности потока носителей заряда граничная концентрация существенно превышает равновесную концентрацию тех же носителей ($p_{гр} \gg p_0$), так как скорость движения носителей ограничена. Следовательно,

$$s_p = \frac{\Phi_p}{p_{гр} - p_0} = \frac{p_{гр} v_p}{p_{гр} - p_0} \approx v_p.$$

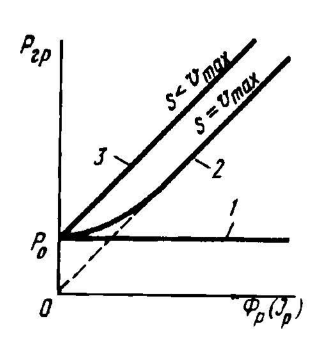


Рис. 2.19. Зависимость граничной концентрации носителей заряда около омического перехода от потока или от плотности тока этих носителей

Таким образом, скорость рекомбинации на омическом переходе не превосходит скорости движения носителей заряда. Если перенос носителей заряда к переходу вызван диффузией, то скорость движения не может превзойти тепловой скорости. Если перенос носителей вызван дрейфом, то скорость тоже ограничена значением максимальной скорости, имеющей порядок тепловой скорости (см. § 1.10).

Итак, скорость рекомбинации на омическом переходе не превосходит тепловой скорости движения носителей заряда. Это имеет практическое значение, так как нет смысла пытаться совершенствовать омические переходы, скорость рекомбинации на которых приближается к максимально возможной.

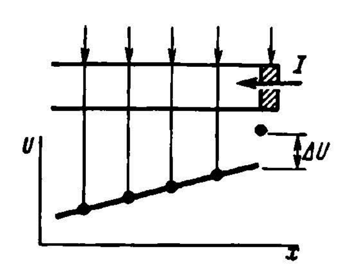


Рис. 2.20. Определение падения напряжения на переходе путем экстраполяции распределения потенциала в полупроводнике

На рис. 2.19 показаны зависимости граничной концентрации носителей заряда около идеального омического перехода с бесконечно большой скоростью рекомбинации (*1*), для реального омического перехода со скоростью рекомбинации, равной максимально возможной скорости движения носителей (*2*), и для реального омического перехода со скоростью рекомбинации, не превышающей максимальной скорости движения носителей (*3*).

Сопротивление омического перехода. Сопротивление омического перехода — это отношение падения напряжения на переходе к проходящему через него току:

$$R_{ом.пер} = \Delta U / I.$$

Чем меньше сопротивление омического перехода, тем он лучше. Для выяснеиия смысла падения иапряжения на омическом переходе проделаем мыслеино следующий опыт. Пусть через омический переход между полупроводником в виде призмы или цилиндра и металлическим электродом проходит ток (рис. 2.20). Потенциал в полупроводиике должен распределиться линейио. В этом можно убедиться, выполнив измерения с помощью зондов. Если одии из зондов поставить на металлический электрод, то полученное значение потенциала не попадает на линейиую зависимость. Под падением иапряжеиия на переходе следует понимать отклонение напряжения иа металлическом электроде от значеиия, полученного путем экстраполяции распределения потенциала в полупроводнике. При практическом определении падения напряжения на переходе не обязательно иметь полупроводник в форме призмы или цилиндра, необходимо только устаиовить реальное распределение потенциала в нем для возможности экстраполяции.

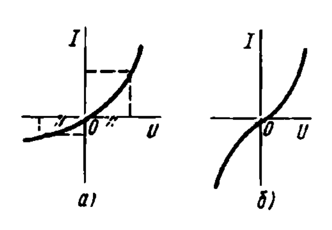


Рис. 2.21. ВАХ омическнх переходов:

а — омический переход, обладающнй некоторым эффектом выпрямления; *б* — омическнй переход с сопротнвленнем, завнсящим от приложенного напряжеиня

Сопротивление омического перехода зависит от его площади. Поэтому можно сравнивать переходы с одинаковой площадью или ввести *удельное сопротивление омического перехода*, определнв его как отношение падения напряжения на переходе к плотности проходящего тока:

$$\varrho_{\text{ом.пер}} = \Delta U / J. \tag{2.44}$$

Размерность удельного сопротивления омического перехода — Ом·см^2.

Сопротнвление перехода связано с удельным сопротивлением перехода:

$$R_{\text{ом.пер}} = \varrho_{\text{ом.пер}} / S.$$

Параметры линейности. В зависимости от того, какие требования предъявляются к линейности вольт-амперной характеристики (ВАХ) омического перехода и каков вид нелииейиости, ее можно характеризовать и оценивать по-разному. Если реальный омический переход обладает эффектом выпрямления (рис. 2.21, *а*), то нелинейиость ВАХ можно оценивать *коэффициентом выпрямления*, т. е. отношением прямого тока к обратному при равных значениях приложенного прямого и обратного напряжений. Идеальный омический переход должен иметь коэффициент выпрямления, равный единице.

Нелинейность симметричной ВАХ омического перехода (рис. 2.21, *б*) можно оценивать *коэффициентом нелинейности* — отношением статического сопротивления к дифференциальному при заданном значении постоянной составляющей тока через переход. Идеальный омический переход с линейной ВАХ будет иметь коэффициент нелинейности, равный единице.

Структура реального невыпрямляющего контакта

Структура реального невыпрямляющего контакта в полупроводниковых приборах, удовлетворяющая перечисленным требованиям, имеет сравнительно сложное строение и состоит из нескольких последовательно соединенных омических переходов (рис. 2.22).

Для меньшей вероятности накопления неосновных носителей заряда около омического перехода между металлом и полупроводником высота потенциального барьера для неосновных носителей заряда должна быть здесь по возможности меньшей. Поэтому металл желательно подобрать с работой выхода, которая мало бы отличалась от работы выхода электронов в полупроводнике (см. § 2.10). Так как это обеспечить трудно, то поверхностный слой полупроводника должен быть сильно легирован соответствующей примесью для обеспечения возможности туннелирования носителей заряда сквозь тонкий потенциальный барьер.

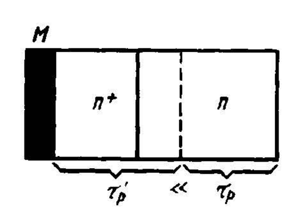


Рис. 2.22. Структура реального невыпрямляющего контакта с последовательно соединенными омическими переходами

Вблизи омического перехода между полупроводниками с одной электропроводностью, но с разной степенью легирования, также может происходить накопление неосновных носителей заряда (см. § 2.9). Для меньшего влияния этого эффекта на характеристики и параметры полупроводникового прибора целесообразно в поверхностный слой полупроводника ввести примеси рекомбинационных ловушек (например, золото), что уменьшит время жизни носителей заряда в этой части структуры. При малом времени жизни накопленные носители заряда будут быстро рекомбинировать, т. е. процесс их рассасывания меньше будет сказываться на характеристиках и параметрах прибора.

Однако в реальных невыпрямляющих контактах часто образуются различные промежуточные слои, ухудшающие свойства омических переходов. Поэтому окончательную доработку технологии их изготовления проводят экспериментально.

Контрольные вопросы

1. Что такое электронно-дырочный переход?

2. Как и почему изменяется высота потенциального барьера *p-n*-перехода с изменением температуры и концентрации примесей в прилегающих к переходу областях?

3. Что такое инжекция и экстракция неосновных носителей заряда?

4. От чего зависит и чем определяется концентрация неосновных носителей заряда на границах *p-n*-перехода?

5. Как можно записать условие электрической нейтральности *p-n*-перехода и каков физический смысл этого условия?

6. Как распределяются напряженность электрического поля и потенциал в резком и в плавном *p-n*-переходах?

7 Как изменяется толщина резкого и плавного *p-n*-переходов при изменении приложенного напряжения?

8. Что такое барьерная емкость *p-n*-перехода?

9. Как определить контактную разность потенциалов на *p-n*-переходе с помощью экспериментальных вольт-фарадных характеристик этого перехода?

10. Почему электрический переход между двумя одинаковыми полупроводниками с одним типом электропроводности, но с разной концентрацией примесей, является омическим и неинжектирующим неосновные носители заряда в высокоомную область?

11. При каких условиях электрический переход между металлом и полупроводником будет омическим?

12. При каких условиях электрический переход между металлом и полупроводником будет выпрямляющим без инжекции неосновных носителей заряда в полупроводник?

13. В каком случае может происходить накопление неосновных носителей заряда вблизи омического перехода между металлом и полупроводником?

14. Каковы правила построения энергетических диаграмм гетеропереходов?

15. Почему и при каких условиях на гетеропереходах может происходить выпрямление без инжекции неосновных носителей заряда?

16. Почему на гетеропереходе между двумя полупроводниками с одним типом электропроводности может наблюдаться эффект выпрямления?

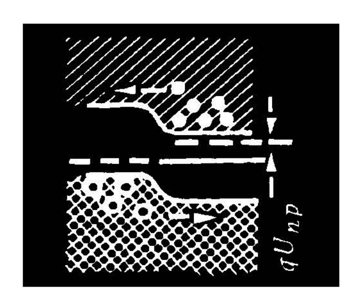
17. Каким требованиям должны удовлетворять омические переходы?

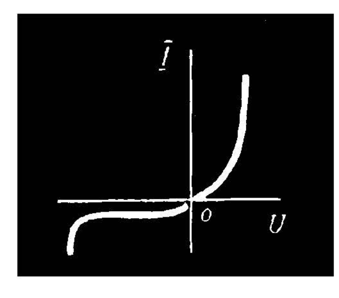
18. Что такое скорость рекомбинации на омическом переходе?

19. Что такое сопротивление омического перехода и как его определить экспериментально?

20. Что такое коэффициент выпрямления и почему этот параметр применим только при наличии слабого эффекта выпрямления?

21. Что такое коэффициент нелинейности ВАХ омического перехода?





глава

3

Полупроводниковые диоды

§ 3.1. СТРУКТУРА И ОСНОВНЫЕ ЭЛЕМЕНТЫ

Полупроводниковый диод — это полупроводниковый прибор с одним выпрямляющим электрическим переходом и двумя выводами, в котором используется то или иное свойство выпрямляющего электрического перехода.

В качестве выпрямляющего электрического перехода в полупроводниковых диодах может быть электронно-дырочный переход, гетеропереход или выпрямляющий переход, образованный в результате контакта между металлом и полупроводником (переход Шотки).

В диоде с *p-n*-переходом или с гетеропереходом кроме выпрямляющего перехода должно быть два омических перехода, через которые *p-* и *n*-области диода соединены с выводами (рис. 3.1, *а*). В диоде с выпрямляющим электрическим переходом в виде контакта между металлом и полупроводником всего один омический переход (рис. 3.1, *б*).

Обычно полупроводниковые диоды имеют несимметричные *p-n*-переходы. Поэтому при полярности внешнего напряжения, при которой происходит понижение потенциального барьера в *p-n*-переходе, т. е. при *прямом направлении для p-n-перехода,* количество носителей заряда, инжектированных из сильнолегированной в слаболегированную область, значительно больше, чем количество носителей, проходящих в противоположном направлении. В соответствии с общим определением (см. § 2.2) область полупроводникового диода, в которую происходит инжекция неосновных для этой области носителей заряда, называют *базой диода*. Итак, в диоде базовой областью является слаболегированная область.

Если к диоду с несимметричным *p-n*-переходом приложено

напряжение, при котором происходит повышение потенциального барьера в p-n-переходе, т. е. в *обратном направлении для p-n-перехода,* то экстракция неосновных носителей заряда будет происходить в основном из базы диода. Таким образом, база диода может оказывать существенное влияние на характеристики и параметры диода.

В зависимости от соотношения линейных размеров выпрямляющего электрического перехода и характеристической длины различают плоскостные и точечные диоды. Характеристической длиной для диода является наименьшая по значению из двух величин, определяющая свойства и характеристики диода: диф-

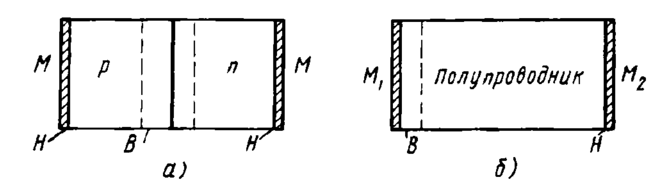


Рис. 3.1. Структуры полупроводниковых диодов:
а — с выпрямляющим электрическим переходом в виде p-n-перехода; *б* — с выпрямляющим электрическим переходом на контакте между металлом и полупроводником; *В* — выпрямляющие электрические переходы; *Н* — невыпрямляющие, т. е. омические, переходы

фузионная длина неосновных носителей заряда в базе или толщина базы.

Плоскостным называют диод, у которого линейные размеры, определяющие площадь выпрямляющего электрического перехода, значительно больше характеристической длины.

Точечным называют диод, у которого линейные размеры, определяющие площадь выпрямляющего электрического перехода, значительно меньше характеристической длины.

В выпрямляющем электрическом переходе и прилегающих к нему областях происходят разнообразные физические процессы, которые могут приводить к эффекту выпрямления, к нелинейному росту тока с увеличением напряжения, к лавинному размножению носителей заряда при ударной ионизации атомов полупроводника, к туннелированию носителей сквозь потенциальный барьер выпрямляющего электрического перехода как при обратном, так в определенных условиях и при прямом напряжении, к изменению барьерной емкости с изменением напряжения, к эффекту накопления и рассасывания неосновных носителей заряда в прилегающих к выпрямляющему переходу областях. Все эти эффекты используют для создания различных видов полупроводниковых диодов: выпрямительных, смесительных, детекторных и переключательных, диодов с резким восстановлением обратного сопротивления, стабилитронов, стабисторов, шумовых, лавинно-пролетных, туннельных и обращенных диодов, варикапов.

Некоторые из перечисленных эффектов являются нежелательными и даже вредными в одних диодах, но в других диодах эти же эффекты могут служить основой принципа действия.

§ 3.2. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА ДИОДА ПРИ ИНЖЕКЦИИ И ЭКСТРАКЦИИ НОСИТЕЛЕЙ ЗАРЯДА

Прямое включение диода

При прямом напряжении на диоде внешнее напряжение частично компенсирует контактную разность потенциалов на *p-n*-переходе, так как внешнее электрическое поле при прямом включении диода направлено противоположно диффузионному полю. Поэтому высота потенциального барьера перехода уменьшается пропорционально приложенному к диоду напряжению. Пренебрегая падением напряжения на базе диода, рассмотрим диод при малых прямых токах.

С уменьшением высоты потенциального барьера увеличивается количество носителей заряда, которые могут преодолеть потенциальный барьер и перейти в соседнюю область диода, где они окажутся неосновными носителями (см. рис. 2.1). Этот процесс, как было отмечено в § 2.2, называют инжекцией неосновных носителей заряда через *p-n*-переход.

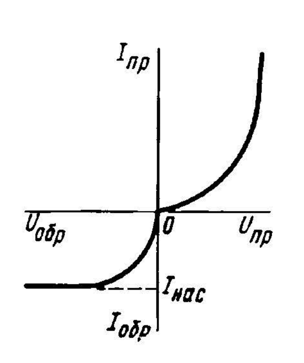


Рис. 3.2. Вольт-амперная характеристика диода при инжекции и экстракции носителей заряда

Так как высота потенциального барьера уменьшается пропорционально приложенному напряжению, а носители заряда распределены по энергиям по экспоненциальному закону в соответствии со статистикой Ферми — Дирака или Максвелла — Больцмана, то прямая ветвь ВАХ диода должна быть похожа на экспоненту (рис. 3.2).

Рассмотрим влияние некоторых факторов на прямую ветвь вольт-амперной характеристики диода.

При увеличении температуры диода уменьшается высота потенциального барьера (см. § 2.1) и изменяется распределение носителей заряда по энергиям (электроны, например, занимают более высокие энергетические уровни в зоне проводимости). Из-за этих двух причин прямой ток через диод увеличивается с ростом температуры при неизменном прямом напряжении (рис. 3.3, *а*).

Если сравнить прямые ветви двух диодов, изготовленных из разных материалов, с разной шириной запрещенной зоны, то у диода из материала с большей шириной запрещенной зоны будет

больше высота потенциальиого барьера (см. § 2.1). Следовательно, прямой ток через диод из материала с большей шириной запрещенной зоны будет меньше при том же прямом иапряжеиии (рис. 3.3, *б*).

С увеличеиием концентрации примесей в прилегающих к *p*-*n*-переходу областях будет увеличиваться высота потенциального барьера перехода (см. § 2.1), а значит, будет меньше прямой ток при том же прямом напряжении (рис. 3.3, *в*).

Обратное включение диода

При обратном включении диода виешнее электрическое поле и диффузиоиное поле в *p*-*n*-переходе совпадают по направлению, происходит экстракция неосновных носителей заряда из прилегающих к переходу областей (см. § 2.2 и рис. 2.1, *е*). Это приводит к уменьшению граничной концентрации неосновных носителей заряда около *p*-*n*-перехода и к появлеиию диффузии неосновных носителей к переходу — идет диффузионный ток неосновных носителей, возникающих в результате тепловой генерации в объеме *n*- и *p*-областей диода, а также на омических переходах.

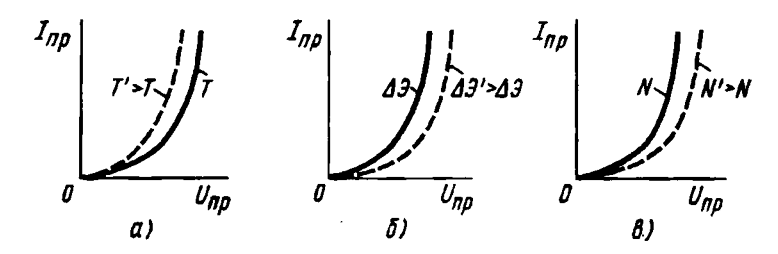


Рис. 3.3. Прямые ветви ВАХ диода при разных температурах (*а*), при разной ширине запрещенной зоны исходного материала (*б*) и при разной концентрации примесей в прилегающих к *p*-*n*-переходу областях (*в*)

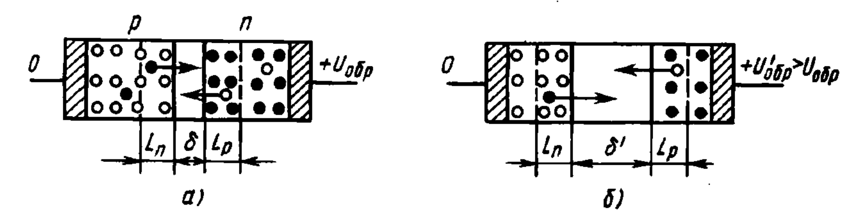


Рис. 3.4. Экстракция неосновных носителей из прилегающих к *p*-*n*-переходу областей при разных обратных напряженнях на диоде

За время жизни до *p*-*n*-перехода могут продиффундировать неосновные носители, возникшие в *n*- и *p*-областях на расстоя-

ии, не превышающем соответствующей диффузионной длины (рис. 3.4, *а, б*). Остальные неосновные носители, не успев дойти до перехода, рекомбинируют в объеме. Это справедливо для разных обратных напряжений на диоде, если толщины прилегающих к переходу областей превышают диффузионные длины неосновных носителей заряда. Поэтому обратный ток начиная с очень малых значений обратного напряжения не будет изменяться с изменением напряжения (см. рис. 3.2). Этот неизменный с изменением напряжения обратный ток через диод называют *током насыщения*.

Рассмотрев физические процессы в диоде при обратном напряжении, можно выразить плотность тока насыщения через параметры полупроводникового материала. Для этого надо вспомнить общее соотношение для плотности тока при наличии носителей заряда двух типов:

$$J = q(pv_p + nv_n),$$

где v_p и v_n — скорости либо диффузии, либо дрейфа дырок и электронов.

В диоде неосновные носители заряда диффундируют к переходу, поэтому их скорости можно представить как диффузионные длины, деленные на соответствующие времена жизни. Вместо общих концентраций p и n подставим концентрации неосновных носителей, так как с движением именно неосновных носителей заряда связан обратный ток диода. Тогда

$$J_{\text{нас}} = q\left(p_{n0}\frac{L_p}{\tau_p} + n_{p0}\frac{L_n}{\tau_n}\right). \qquad (3.1)$$

Если учесть соотношения (1.19), а также практически полную ионизацию примесей при комнатных температурах, то (3.1) можно привести к виду

$$J_{\text{нас}} = qn_i^2\left(\frac{L_p}{N_{\text{д}n}\tau_p} + \frac{L_n}{N_{\text{а}p}\tau_n}\right). \qquad (3.2)$$

При увеличении температуры диода плотность тока насыщения увеличивается, так как с температурой экспоненциально растет собственная концентрация носителей заряда [см. (1.15)].

В диодах на основе материала с большей шириной запрещенной зоны плотность тока насыщения должна быть значительно меньше, так как собственная концентрация экспоненциально уменьшается с увеличением ширины запрещенной зоны [см. (1.15)]. Сравнивая германиевые и кремниевые диоды и учитывая разницу в собственных концентрациях носителей в германии и кремнии, которая составляет три порядка (см. § 1.4), следует заключить, что плотность тока насыщения в кремниевых диодах должна быть меньше на шесть порядков.

С увеличением концентрации примесей в прилегающих к переходу областях плотность тока насыщения в соответствии с (3.2) должна уменьшиться.

§ 3.3. РАСЧЕТ РАСПРЕДЕЛЕНИЯ НЕОСНОВНЫХ НОСИТЕЛЕЙ ЗАРЯДА В БАЗЕ ДИОДА

В качестве примера проведем расчет для диода с иесимметричным электронио-дырочиым переходом p^+-n (рис. 3.5) при приложеиии к нему напряжения, имеющего постоянную и малую переменную составляющие:

$$u = U + \dot{U}_m \exp j\omega t. \qquad (3.3)$$

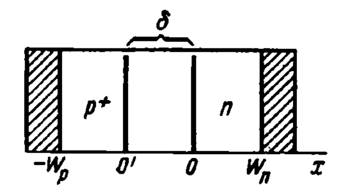


Рис. 3.5. Одномерная модель диода, принятая для расчета

Здесь перемеиная составляющая записана в виде комплексной величины в показательной форме. Так как в (3.3) складываются постояиная и переменная составляющие, то физический смысл имеет только проекция вектора переменного напряжения на действительную ось.

Условием малости переменной составляющей напряжения для удобства математических преобразований выберем следующее:

$$U_m \ll kT/q, \qquad (3.4)$$

т. е. амплитуда переменной составляющей напряжения не должна превосходить ~ 10 мВ.

Допущения при расчете

Для облегчеиия расчетов обычно выбирают упрощенную модель структуры того или иного прибора. В данном случае допустим, что:

1) p-n-переход диода плоский, т. е. будем рассматривать одиомерную модель диода (рис. 3.5);

2) токи малы и не вызывают существенного падения напряжения иа сопротивлении базы диода; таким образом, электрическое поле сосредоточено только в p- n-переходе;

3) омические переходы идеальны, т. е. около них в полупроводнике существует всегда только равновесная концентрация иосителей заряда;

4) поверхностные явления несущественны;

5) в p-n-переходе не происходят процессы генерации нли рекомбинации иеравновесиых носителей заряда;

6) рекомбинация неосновных носителей в объеме базы линейная, т. е. количество носителей, рекомбинирующих в единице объема за единицу времени, прямо пропорционально избыточной концентрации:

$$R_p = \frac{\Delta p_n}{\tau_p} = \frac{p_n - p_{n0}}{\tau_p}.$$

Дифференциальные уравнения

Основным для решеиия поставленной задачи является *уравнение непрерывности,* например, для дырок в базе диода с электропроводиостью n-типа

$$\frac{\partial p_n}{\partial t} = -\frac{1}{q}\operatorname{div}\mathbf{J}_p - R_p + G_p. \tag{3.5}$$

Это уравиеиие показывает, как и по каким причинам изменяется концентрация дырок со временем. Во-первых, концеитрация дырок может изменяться из-за существоваиия дивергенции тока дырок, что учитывает первое слагаемое. Во-вторых, коицентрация дырок может измеияться из-за их рекомбинации, что учитывает второе слагаемое (R_p — скорость рекомбинации). Это же слагаемое в зависимости от знака может учитывать измеиеиие концеитрации дырок из-за тепловой генерации. В-третьих, концентрация дырок может измеияться из-за нетепловой генерации (ударная ионизация, иоиизация под действием света и т. д.). В даниом случае $G_p = 0$.

С учетом прииятых в начале параграфа допущений уравнение непрерывности (3.5) перепишем следующим образом:

$$\frac{\partial p_n}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x} - \frac{p_n - p_{n0}}{\tau_p}. \tag{3.6}$$

Воспользуемся уравиением (1.31) для плотиости тока дырок, также упростив его с учетом прииятых допущеиий:

$$\mathbf{J}_p = -qD_p\frac{\partial p_n}{\partial x}. \tag{3.7}$$

После подстановки (3.7) в (3.6) имеем

$$\frac{\partial p_n}{\partial t} = D_p\frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{\tau_p}, \tag{3.8}$$

т. е. получеио дифференциальное уравнение второго порядка в частиых производных. Для его решения необходимы граиичные и начальиые условия.

Граничные условия

При малых токах концентрация иеосиовных носителей заряда в базе около p-n-перехода определяется соотиошением (2.5):

$$p_n(0) = p_{n0}\exp\frac{qu}{kT}.$$

Подставив сюда значение напряжения (3.3), получим при $x = 0$

$$p_n(0) = p_{n0}\exp\frac{q}{kT}(U + \dot{U}_m\exp j\omega t) =$$

$$= p_{n0}\left(\exp\frac{qU}{kT}\right)\exp\left(\frac{q\dot{U}_m}{kT}\exp j\omega t\right). \tag{3.9}$$

Аргумент второй экспоненты мал. Поэтому ее можно разложить в ряд, ограничившись двумя первыми его членами:

$$\exp y \approx 1 + y + \ldots .$$

Тогда (3.9) с учетом разложения в ряд примет вид

$$p_n(0) = p_{n0} \exp \frac{qU}{kT} + p_{n0}\left(\exp \frac{qU}{kT}\right)\frac{q\dot{U}_m}{kT} \exp j\omega t. \qquad (3.10)$$

Таким образом, концентрация неосновных носителей заряда в базе около *p-n*-перехода имеет постоянную и переменную составляющие. Частота изменения переменной составляющей та же, что и частота приложенного переменного напряжения. Если бы переменное напряжение не было мало, то у переменной составляющей концентрации носителей появились бы гармоники высшего порядка с частотой, кратной ω.

Второе граничное условие следует из идеальности омического перехода, т. е. при $x = W_n$

$$p_n(W_n) = p_{n0}. \qquad (3.11)$$

Форма решения

Для преобразования дифференциального уравнения (3.8) выберем форму решения в виде суммы постоянной и переменной составляющих концентрации, т. е. форму, аналогичную форме граничного условия (3.10). Это обусловлено линейностью уравнения для концентрации носителей заряда, так что новые гармоники появиться не могут. Тогда решение дифференциального уравнения (3.8) должно иметь вид

$$p_n(x,\ t) = p_{n0} + \Delta p_n(x) + \dot{p}_n(x) \exp j\omega t, \qquad (3.12)$$

где $p_n(x;\ t)$ — полная концентрация неосновных носителей заряда в базе; $\Delta p_n(x)$ — постоянная составляющая избыточной концентрации неосновных носителей, зависящая только от координаты; $\dot{p}_n(x) \exp j\omega t$ — переменная составляющая избыточной концентрации неосновных носителей заряда, зависящая как от координаты, так и от времени.

Таким образом, постоянная составляющая в (3.12) представлена в виде суммы равновесной и избыточной концентраций.

Преобразование общего дифференциального уравнения

После подстановки выбранной формы решения (3.12) в (3.8) получим

$$\dot{p}_n(x) j\omega \exp j\omega t = D_p \frac{\partial^2[\Delta p_n(x)]}{\partial x^2} + D_p \frac{\partial^2 \dot{p}_n(x)}{\partial x^2} \exp j\omega t -$$

$$- \frac{\Delta p_n(x)}{\tau_p} - \frac{\dot{p}_n(x)}{\tau_p} \exp j\omega t. \qquad (3.13)$$

В этом уравнении есть слагаемые, зависящие и не зависящие от времени. Уравнение справедливо только в том случае, если алгебраические суммы не зависящих и зависящих от времени составляющих отдельно равны нулю. Поэтому для постоянной составляющей избыточной концентрации из (3.13) получим

$$\frac{d^2[\Delta p_n(x)]}{d^2x} = \frac{\Delta p_n(x)}{L_p^2}, \tag{3.14}$$

где

$$L_p = \sqrt{D_p \tau_p}.$$

Граничные условия для постоянной составляющей избыточной концентрации:

при $x = 0$ $$\Delta p_n(0) = p_{n0}(\exp \frac{qU}{kT} - 1), \tag{3.15}$$

при $x = W_n$ $$\Delta p_n(W_n) = 0. \tag{3.16}$$

Для переменной составляющей избыточной концентрации

$$\frac{d^2 \dot{p}_n(x)}{dx^2} = \frac{\dot{p}_n(x)}{\Lambda_p^2}, \tag{3.17}$$

где

$$\Lambda_p = \frac{L_p}{\sqrt{1 + j\omega\tau_p}}.$$

Граничные условия для переменной составляющей избыточной концентрации запишем, исходя из общих граничных условий (3.10) и (3.11):

при $x = 0$ $$\dot{p}_n(0)\exp j\omega t = p_{n0}\left(\exp \frac{qU}{kT}\right)\frac{q\dot{U}_m}{kT}\exp j\omega t,$$

или окончательно

$$\dot{p}_n(0) = p_{n0}\left(\exp \frac{qU}{kT}\right)\frac{q\dot{U}_m}{kT}; \tag{3.18}$$

при $x = W_n$ $$\dot{p}_n(W_n) = 0. \tag{3.19}$$

Полученные уравнения для постоянной и переменной составляющих аналогичны, так что решать можно только одно из них.

Решение для постоянной составляющей избыточной концентрации неосновных носителей в базе

Решение дифференциального уравнения (3.14) удобно искать в виде

$$\Delta p_n(x) = A_1 \mathrm{ch} \frac{x}{L_p} + A_2 \mathrm{sh} \frac{x}{L_p}. \tag{3.20}$$

Такой вид решения упрощает поиск произвольных постоянных, если заданы условия на границах и при этом на одной границе — нуль.

Подставив в (3.20) $x = 0$ и учтя граничное условие (3.15), получим

$$\Delta p_n(0) = p_{n0}\left(\exp\frac{qU}{kT} - 1\right) = A_1.$$

С учетом граничного условия (3.16) и значения постоянной интегрирования A_1 при $x = W_n$

$$\Delta p_n(W_n) = 0 = p_{n0}\left(\exp\frac{qU}{kT} - 1\right)\operatorname{ch}\frac{W_n}{L_p} + A_2\operatorname{sh}\frac{W_n}{L_p}.$$

Следовательно,

$$A_2 = -p_{n0}\left(\exp\frac{qU}{kT} - 1\right)\operatorname{cth}\frac{W_n}{L_p}.$$

Подставив в (3.20) значения постоянных интегрирования, окончательно получим

$$\Delta p_n(x) = p_{n0}\left(\exp\frac{qU}{kT} - 1\right)\left(\operatorname{ch}\frac{x}{L_p} - \operatorname{cth}\frac{W_n}{L_p}\operatorname{sh}\frac{x}{L_p}\right). \qquad (3.21)$$

Таково распределение постоянной составляющей избыточной концентрации неосновных носителей заряда в базе диода при разных напряжениях.

§ 3.4. РАСЧЕТ ПОСТОЯННЫХ ТОКОВ, ПРОХОДЯЩИХ ЧЕРЕЗ ДИОД И СВЯЗАННЫХ С ИНЖЕКЦИЕЙ И ЭКСТРАКЦИЕЙ НОСИТЕЛЕЙ ЗАРЯДА

Для определения дырочной составляющей плотности тока в произвольном сечении базы диода воспользуемся соотношениями (3.7) и (3.21). После дифференцирования получим

$$J_p(x) = -\frac{qD_p p_{n0}}{L_p}\left(\operatorname{sh}\frac{x}{L_p} - \operatorname{cth}\frac{W_n}{L_p}\operatorname{ch}\frac{x}{L_p}\right)\left(\exp\frac{qU}{kT} - 1\right). \qquad (3.22)$$

Такова плотность постоянного тока в различных частях базы диода.

При $x = 0$, т. е. на границе базы с p-n-переходом, дырочная составляющая плотности тока через переход

$$J_p(0) = \frac{qD_p p_{n0}}{L_p}\operatorname{cth}\frac{W_n}{L_p}\left(\exp\frac{qU}{kT} - 1\right). \qquad (3.23)$$

Аналогично можно записать и электронную составляющую плотности тока через переход диода, т. е. через границу между p-областью и p-n-переходом:

$$J_n(0') = \frac{qD_n n_{p0}}{L_n} \operatorname{cth} \frac{W_p}{L_n} \left(\exp \frac{qU}{kT} - 1\right). \quad (3.24)$$

Для практики больший интерес представляет знание полной плотности тока, т. е. суммы электронной и дырочной составляющих. Суммировать надо плотности токов в одном и том же сечении. Однако расчет тока основных носителей заряда в базе потребовал бы использования другой методики по сравнению с методикой расчета тока неосновных носителей. Поэтому используем допущение, что в p-n-переходе не происходят процессы генерации и рекомбинации носителей. Следовательно, составляющие плотностей тока (электронного и дырочного) одинаковы по обе стороны перехода. Теперь можно определить плотности тока неосновных носителей заряда с двух сторон p-n-перехода и сложить их:

$$J = J_p(0) + J_n(0') = q\left(\frac{p_{n0}D_p}{L_p} \operatorname{cth} \frac{W_n}{L_p} + \frac{n_{p0}D_n}{L_n} \operatorname{cth} \frac{W_p}{L_n}\right) \times$$
$$\times \left(\exp \frac{qU}{kT} - 1\right). \quad (3.25)$$

Значение второго сомножителя в (3.25) определяется параметрами полупроводниковых областей с двух сторон p-n-перехода и толщиной этих областей. Произведение первых двух сомножителей, не зависящее от напряжения, называют плотностью тока насыщения:

$$J_{нас} = q\left(\frac{p_{n0}D_p}{L_p} \operatorname{cth} \frac{W_n}{L_p} + \frac{n_{p0}D_n}{L_n} \operatorname{cth} \frac{W_p}{L_n}\right). \quad (3.26)$$

Строго говоря, этот ток не совсем и не всегда насыщенный, так как толщина базы диода зависит от напряжения в связи с изменением толщины p-n-перехода при изменении приложенного напряжения.

Таким образом, ВАХ диода обычно записывают в виде

$$J = J_{нас}\left(\exp \frac{qU}{kT} - 1\right). \quad (3.27)$$

Графическое изображение ВАХ показано на рис. 3.2. Для удобства масштабы прямых и обратных напряжений, а также прямых и обратных токов выбирают разными.

§ 3.5. ЧАСТНЫЕ СЛУЧАИ РАСЧЕТА РАСПРЕДЕЛЕНИЯ НЕОСНОВНЫХ НОСИТЕЛЕЙ ЗАРЯДА И ТОКА НАСЫЩЕНИЯ

Чтобы получить еще более простые выражения и лучше разобраться в смысле результата, рассмотрим предельные частные случаи диодов с толстой и тонкой базой.

Диод с толстой базой — это диод, толщина базы которого значительно превышает диффузионную длину неосновных носи-

телей заряда ($W_n \gg L_p$). Для другой области диода из-за сильного ее легирования тем более справедливо подобное неравенство, т. е. $W_p \gg L_n$. Тогда аргументы гиперболических котангенсов в соотношении (3.21) значительно превышают единицу, а сами гиперболические котангенсы близки к единице:

$$\mathrm{cth}\frac{W_n}{L_p} \approx 1 \text{ и } \mathrm{cth}\frac{W_p}{L_n} \approx 1.$$

Воспользовавшись формулами Эйлера $\left(\mathrm{sh}\,y = \frac{e^y - e^{-y}}{2}\ \ \mathrm{ch}\,y = \frac{e^y + e^{-y}}{2}\right)$, из соотношения (3.21) получим распределение избыточной концентрации неосновных носителей заряда в базе диода с толстой базой:

$$\Delta p_n(x) = p_{n0}\left(\exp\frac{qU}{kT} - 1\right)\exp\left(-\frac{x}{L_p}\right). \quad (3.28)$$

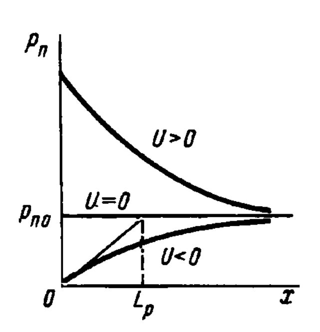


Рис. 3.6. Распределение концентрации неосновных носителей заряда в базе диода с толстой базой при разных напряжениях

Следовательно, в диоде с толстой базой абсолютное значение избыточной концентрации неосновных носителей в базе ($\Delta p_n = p_n - p_{n0}$) экспоненциально уменьшается с увеличением расстояния от p-n-перехода (рис. 3.6). При прямых напряжениях избыточная концентрация положительна, что соответствует инжекции неосновных носителей в базу. При обратных напряжениях избыточная концентрация отрицательна, т. е. $p_n < p_{n0}$, что соответствует экстракции неосновных носителей из базы.

Выражение для плотности тока насыщения в диоде с толстой базой можно найти из соотношения (3.26) с учетом того, что $\mathrm{cth}\frac{W}{L} \approx 1$:

$$J_{\text{нас}} = q\left(\frac{p_{n0}D_p}{L_p} + \frac{n_{p0}D_n}{L_n}\right). \quad (3.29)$$

Это выражение для плотности тока насыщения в диоде с толстой базой можно получить также, исходя из того, что обратный ток обусловлен только диффузией неосновных носителей заряда в прилегающих к p-n-переходу областях. Поэтому для вычисления плотности тока насыщения надо воспользоваться вторыми слагаемыми в (1.30) и (1.31) или соотношениями (1.26) и (1.27). При этом градиенты концентраций неосновных носителей в n- и p-областях около p-n-перехода можно определить как p_{n0}/L_p и n_{p0}/L_n (рис. 3.6).

Таким образом, в диоде с толстой базой плотность тока насыщения не зависит от напряжения. Поэтому обратный ток через диод с учетом только экстракции неосновных носителей заряда начиная с обратных напряжений $|U_{обр}| \gg kT/q$ или $|U_{обр}| > 0,1$ В, не изменяется с напряжением (рис. 3.7). Тот же вывод можно сделать, если представить распределение неосновных носителей в базе диода с толстой базой при разных обратных напряжениях (рис. 3.8). С увеличением обратного напряжения по абсолютному значению и соответственно с увеличением толщины *p-n*-перехода

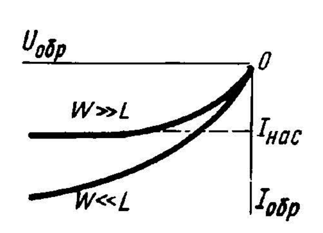


Рис. 3.7. Обратные ветви ВАХ диодов с толстой и с тонкой базой при учете экстракции неосновных носителей заряда из прилегающих к *p-n*-переходу областей

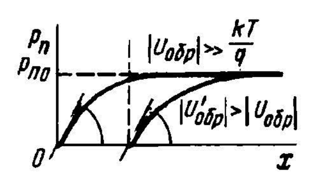


Рис. 3.8. Распределение концентрации неосновных носителей заряда в базе диода при разных обратных напряжениях, поясняющее неизменность обратного тока (тока насыщения) в диоде с толстой базой при экстракции неосновных носителей

за счет толщины базы происходит смещение кривых распределения концентрации неосновных носителей заряда в глубь базы при неизменном градиенте концентрации около *p-n*-перехода, что по соотношению (1.27) соответствует неизменному току.

Выражение (3.29) аналогично выражению (3.1), полученному из чисто физических соображений, так как $L = \sqrt{D\tau}$.

В соответствии с (3.29) характеристической длиной, определяющей свойства и многие параметры диода с толстой базой, является диффузионная длина неосновных носителей заряда в базе диода.

Диод с тонкой базой — это диод, толщина базы которого значительно меньше диффузионной длины неосновных носителей заряда ($W_n \ll L_p$). В этом случае аргументы всех гиперболических функций в соотношении (3.21) будут малыми (меньше единицы). Поэтому, раскладывая гиперболические функции в ряд, можно ограничиться всего одним членом разложения ($\mathrm{cth}\, y \approx 1/y$; $\mathrm{sh}\, y \approx y$; $\mathrm{ch}\, y \approx 1$). Тогда для распределения концентрации неосновных носителей в диоде с тонкой базой из (3.21) получим

$$\Delta p_n(x) = p_{n0}\left(\exp\frac{qU}{kT} - 1\right)\left(1 - \frac{x}{W_n}\right). \qquad (3.30)$$

Следовательно, в диоде с тонкой базой концентрация неосновных носителей заряда уменьшается по мере удаления от p-n-перехода линейно (рис. 3.9). Другими словами, например, для диода с несимметричным p^+-n-переходом плотность тока в любом сечении базы неизменна в соответствии с (3.7) и с учетом (3.30), т. е. рекомбинация неосновных носителей заряда в базе несущественна. Все неосновные носители заряда, инжектированные в базу при прямом напряжении, доходят до омического перехода, где и рекомбинируют. При обратном напряжении через p-n-переход диода с тонкой базой происходит экстракция всех носителей заряда, поставляемых в базу невыпрямляющим контактом.

Плотность тока насыщения в диоде с тонкой базой из общего соотношения (3.26) с учетом разложения в ряд гиперболических функций

$$J_{нас} = q\left(\frac{p_{n0}D_p}{W_n} + \frac{n_{p0}D_n}{W_p}\right). \quad (3.31)$$

Это выражение для плотности тока насыщения можно также получить, исходя из того, что обратный ток обусловлен только диффузией неосновных носителей заряда от омических переходов к p-n-переходу по прилегающим к переходу областям. Поэтому для вычисления плотности тока насыщения надо воспользоваться вторыми слагаемыми в (1.30) и (1.31) или соотношениями (1.26) и (1.27). При этом градиенты концентраций неосновных носителей в n- и p-областях можно определить как p_{n0}/W_n и n_{p0}/W_p (рис. 3.9).

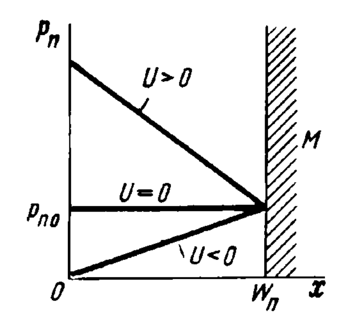


Рис. 3.9. Распределение концентрации неосновных носителей заряда в базе диода с тонкой базой при разных напряжениях

Таким образом, в диоде с тонкой базой плотность тока насыщения зависит от обратного напряжения, так как с изменением обратного напряжения изменяется и толщина базы (W_n или W_p) в связи с изменением толщины p-n-перехода (см. рис. 3.7). Тот же вывод можно сделать, если представить распределение неосновных носителей заряда в базе диода с тонкой базой при разных обратных напряжениях (рис. 3.10). С увеличением обратного напряжения по абсолютному значению происходит изменение градиента концентрации неосновных носителей заряда в базе, т. е. омический переход (источник неос-

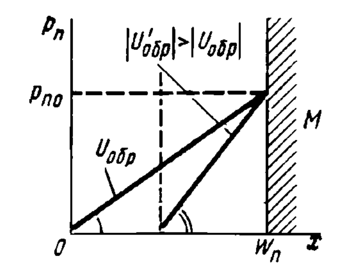


Рис. 3.10. Распределение концентрации неосновных носителей заряда в базе диода при разных обратных напряжениях, поясняющее увеличение обратного тока в диоде с тонкой базой

новных носителей заряда в данном случае) тем сильнее влияет на обратный ток, чем он ближе оказывается к p-n-переходу.

В соответствии с (3.31) характеристической длиной, определяющей свойства и многие параметры диода с тонкой базой, является толщина базы диода.

§ 3.6. РАСЧЕТ ПЕРЕМЕННЫХ ТОКОВ И ПОЛНОЙ ПРОВОДИМОСТИ ДИОДА

В связи с аналогией дифференциальных уравнений для распределения постоянной и переменной составляющих концентрации неосновных носителей (3.14) и (3.17), а также в связи с аналогией граничных условий (3.15) и (3.18), (3.16) и (3.19) можно сразу записать выражение для переменной составляющей плотности тока через диод. Для этого достаточно сделать следующие замены: вместо L подставить

$$\Lambda = \frac{L}{\sqrt{1 + j\omega\tau}}; \qquad (3.32)$$

вместо $\left(\exp\frac{qU}{kT} - 1\right)$ подставить $\left(\exp\frac{qU}{kT}\right)\frac{q\dot{U}_m}{kT}$.

В результате получим выражение для переменной составляющей плотности тока, аналогичное выражению (3.25):

$$\dot{J}_m = \frac{q^2\dot{U}_m}{kT}\left(\frac{p_{n0}D_p}{\Lambda_p}\operatorname{cth}\frac{W_n}{\Lambda_p} + \frac{n_{p0}D_n}{\Lambda_n}\operatorname{cth}\frac{W_p}{\Lambda_n}\right)\exp\frac{qU}{kT}. \qquad (3.33)$$

Видно, что переменная составляющая плотности тока через диод линейно связана с переменным напряжением, если это напряжение мало (см. § 3.3). Отсюда следует, что для описания свойств диода целесообразно воспользоваться обычным приемом электротехники — ввести полную проводимость (либо полное сопротивление) диода для переменного тока:

$$\dot{Y} = \dot{I}_m/\dot{U}_m = S\dot{J}_m/\dot{U}_m.$$

Тогда для диода, используя (3.33), можем записать

$$\dot{Y} = \frac{q^2 S}{kT}\left(\frac{p_{n0}D_p}{\Lambda_p}\operatorname{cth}\frac{W_n}{\Lambda_p} + \frac{n_{p0}D_n}{\Lambda_n}\operatorname{cth}\frac{W_p}{\Lambda_n}\right)\exp\frac{qU}{kT}. \qquad (3.34)$$

Полученное выражение позволяет выполнить расчет проводимости полупроводникового диода на любой частоте и при любых соотношениях размеров диода и диффузионной длины.

Как видно из (3.4), проводимость диода для переменного тока оказывается комплексной. Запишем проводимость диода в виде

$$\dot{Y} = \frac{1}{r} + j\omega C_{\text{диф}},$$

что соответствует параллельной схеме замещения.

Значения *диффузионной емкости* $C_{диф}$ и дифференциального сопротивления r могут быть вычислены для любого случая из общего выражения полной проводимости диода (3.34) путем соответствующих преобразований (извлечения корней из комплексных чисел, взятия гиперболического котангенса сумм и т. д.). Однако получающиеся выражения оказываются довольно сложными, поэтому целесообразно ограничиться частными случаями. Таких частных случаев можно выделить четыре — два соотношения размеров ($W \gg L$ и $W \ll L$) для двух диапазонов частот (высоких и низких).

1) $W \gg L$, низкие частоты. Условие $W \gg L$ означает, что аргументы гиперболических котангенсов в (3.34) являются большими. Следовательно,

$$\operatorname{cth}\left(\frac{W}{L}\sqrt{1+j\omega\tau}\right) \approx 1.$$

Теперь остается преобразовать соотношение (3.32). Преобразование этого соотношения оказывается простым, если $\omega\tau \ll 1$. Такое условие и будем считать критерием низкой частоты. Тогда, раскладывая в биномнальный ряд $\sqrt{1+j\omega\tau}$, получим

$$\frac{1}{\Lambda} \approx \frac{1}{L}\left(1+j\frac{\omega\tau}{2}\right)=\frac{1}{L}+j\omega\frac{\tau}{2L}.$$

Подставив это выражение в формулу для полной проводимости (3.34), после группирования слагаемых получим

$$\dot{Y}=\frac{q^2S}{kT}\left[\left(\frac{p_{n0}D_p}{L_p}+\frac{n_{p0}D_n}{L_n}\right)+\frac{j\omega}{2}(p_{n0}L_p+n_{p0}L_n)\right]\exp\frac{qU}{kT}. \quad (3.35)$$

Следовательно,

$$\frac{1}{r}=\frac{q^2S}{kT}\left(\frac{p_{n0}D_p}{L_p}+\frac{n_{p0}D_n}{L_n}\right)\exp\frac{qU}{kT},$$

или с учетом (3.29)

$$\frac{1}{r}=\frac{qS}{kT}J_{нас}\exp\frac{dU}{kT}.$$

Воспользовавшись выражением (3.27) для ВАХ, запишем

$$I=I_{нас}\left(\exp\frac{qU}{kT}-1\right) \quad \text{или} \quad I_{нас}\exp\frac{qU}{kT}=I+I_{нас}, \quad (3.36)$$

окончательно получим

$$\frac{1}{r}=\frac{q}{kT}(I+I_{нас}). \quad (3.37)$$

Как видно из (3.35), диффузионная емкость

$$C_{диф}=\frac{q^2S}{2kT}(p_{n0}L_p+n_{p0}L_n)\exp\frac{qU}{kT}. \quad (3.38)$$

Если $p_{n0} \gg n_{p0}$, либо $n_{p0} \gg p_{n0}$, либо $\tau_p = \tau_n = \tau$, то выражение для диффузионной емкости с учетом (3.38) и (3.1) или (3.29) приводим к виду

$$C_{\text{диф}} = \frac{qS}{2kT} J_{\text{нас}} \left(\exp \frac{qU}{kT}\right)\tau,$$

или с учетом (3.36)

$$C_{\text{диф}} = \frac{q}{2kT}(I + I_{\text{нас}})\tau. \qquad (3.39)$$

Для диодов с несимметричиым p-n-переходом τ в этой формуле — время жизни неосновных иосителей в слаболегированной области, т. е. в базе диода.

Для такого значения диффузиоииой емкости можно легко найти постоянную времени:

$$rC_{\text{диф}} = \tau/2. \qquad (3.40)$$

2) $W \gg L$, высокие частоты. Критерий высокой частоты $\omega\tau \gg 1$, т. е. нарушение условия низкой частоты. При этом, как и ранее,

$$\operatorname{cth}\left(\frac{W}{L}\sqrt{1 + j\omega\tau}\right) \approx 1.$$

Теперь надо преобразовать (3.32) с учетом высокой частоты. Чтобы извлечь корень, удобно воспользоваться условием $\omega\tau \gg 1$. Тогда, применяя формулу Муавра, получим

$$\sqrt{1 + j\omega\tau} \approx \sqrt{j\omega\tau} = \sqrt{\cos 90^\circ + j\omega\tau \sin 90^\circ} =$$

$$= \sqrt{\omega\tau}(\cos 45^\circ + j \sin 45^\circ) = \sqrt{\omega\tau}\left(\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right),$$

т. е.

$$\frac{1}{\Lambda} \approx \frac{1}{L}\sqrt{\frac{\omega\tau}{2}} + j\frac{1}{L}\sqrt{\frac{\omega\tau}{2}}$$

Поэтому полная проводимость диода

$$\dot{Y} = \frac{q^2 S}{kT}\left[\sqrt{\frac{\omega}{2}}\left(\frac{p_{n0} D_p \sqrt{\tau_p}}{L_p} + \frac{n_{p0} D_n \sqrt{\tau_n}}{L_n}\right) + \right.$$

$$\left. + j\sqrt{\frac{\omega}{2}}\left(\frac{p_{n0} D_p \sqrt{\tau_p}}{L_p} + \frac{n_{p0} D_n \sqrt{\tau_n}}{L_n}\right)\right]\exp\frac{qU}{kT}. \qquad (3.41)$$

Тогда

$$\frac{1}{r} = \frac{q^2 S}{kT}\sqrt{\frac{\omega}{2}}\left(\frac{p_{n0} D_p \sqrt{\tau_p}}{L_p} + \frac{n_{p0} D_n \sqrt{\tau_n}}{L_n}\right)\exp\frac{qU}{kT}. \qquad (3.42)$$

Если $p_{n0} \gg n_{p0}$, либо $n_{p0} \gg p_{n0}$, либо $\tau_p = \tau_n = \tau$, то выражение для активной проводимости

$$\frac{1}{r} = \frac{qS}{kT} \sqrt{\frac{\omega\tau}{2}} J_{нас} \exp \frac{qU}{kT},$$

т. е.

$$\frac{1}{r} = \frac{q}{kT}(I + I_{нас}) \sqrt{\frac{\omega\tau}{2}}. \qquad (3.43)$$

Сравнив (3.43) и (3.37), получим

$$\left(\frac{1}{r}\right)_{вч} = \left(\frac{1}{r}\right)_{нч} \sqrt{\frac{\omega\tau}{2}}. \qquad (3.44)$$

Аналогично, сравнив мнимые части (3.41) и (3.35), имеем

$$(C_{диф})_{вч} = \frac{q}{kT}(I + I_{нас}) \sqrt{\frac{\tau}{2\omega}} = (C_{диф})_{нч} \sqrt{\frac{2}{\omega\tau}}. \qquad (3.45)$$

Постоянная времени в данном случае

$$rC_{диф} = 1/\omega \qquad (3.46)$$

подтверждает, что фазовый сдвиг между током и напряжеиием составляет $\pi/4$. Это же можно заключить из равенства действительной и мнимой частей полной проводимости диода (3.41), т. е.

$$\varphi = \mathrm{arctg}\,\frac{x_c}{r} = \mathrm{arctg}\,1 = 45°.$$

3) **$W \ll L$, низкие частоты.** В данном случае критерий низких частот отличается от того, который применялся для диода с толстой базой. Действительно, условие тонкой базы диода ($W \ll L$) при низких частотах ($\omega \to 0$) приводит к тому, что аргументы гиперболических котангенсов в (3.34) будут малыми. Именно малость аргумента гиперболического котангенса будем считать условием низкой частоты, т. е. критерий низких частот в данном случае

$$\left|\frac{W}{L}\sqrt{1 + j\omega\tau}\right| \ll 1.$$

Воспользовавшись этим условием, разложим гиперболический котангенс в ряд. Если ограничиться одним членом ряда, как это было сделано в § 3.5, то в выражении для проводимости исчезнет мнимая составляющая. Поэтому возьмем два члена ряда, чтобы не потерять интересующий нас эффект. Тогда

$$\mathrm{cth}\,\frac{W}{\Lambda} \approx \frac{\Lambda}{W} + \frac{W}{3\Lambda} = \frac{\Lambda}{W}\left(1 + \frac{W^2}{3\Lambda^2}\right),$$

или

$$\mathrm{cth}\,\frac{W}{\Lambda} \approx \frac{\Lambda}{W}\left(1 + \frac{W^2}{3L^2} + j\omega\tau\frac{W^2}{3L^2}\right) \approx \frac{\Lambda}{W}\left(1 + j\omega\tau\frac{W^2}{3L^2}\right). \qquad (3.47)$$

В полученном выражении слагаемое $W^2/(3L^2)$ пренебрежимо мало по сравнению с единицей. Хотя мнимое слагаемое тоже мало по сравнению с единицей, но пренебрегать им нельзя, опять-таки чтобы не потерять интересующее нас значение мнимой составляющей проводимости. Такой подход приводит к ошибке в действительной составляющей проводимости. Ошибка получается порядка W^2/L^2, что, конечно, несущественно, так как $W \ll L$.

Подставив (3.47) в формулу полной проводимости диода (3.34), после сокращения получим

$$\dot{Y} = \frac{q^2S}{kT}\left[\left(\frac{p_{n0}D_p}{W_n} + \frac{n_{p0}D_n}{W_p}\right) + \frac{j\omega}{3}(p_{n0}W_n + n_{p0}W_p)\right]\exp\frac{qU}{kT}. \qquad (3.48)$$

Тогда с учетом (3.31)

$$\frac{1}{r} = \frac{qS}{kT}J_{\text{нас}}\exp\frac{qU}{kT},$$

или

$$\frac{1}{r} = \frac{q}{kT}(I + I_{\text{нас}}). \qquad (3.49)$$

Диффузионная емкость диода с тонкой базой

$$C_{\text{диф}} = \frac{q^2S}{3kT}(p_{no}W_n + n_{p0}W_p)\exp\frac{qU}{kT}. \qquad (3.50)$$

Если $p_{n0} \gg n_{p0}$ или $n_{p0} \gg p_{n0}$, то диффузионная емкость

$$C_{\text{диф}} = \frac{q}{kT}(I + I_{\text{нас}})\frac{W^2}{3D}, \qquad (3.51)$$

где W и D относятся к слаболегированной области, т. е. к базе диода.

Постоянную времени в данном случае выразим, используя формулы (3.50) и (3.51):

$$rC_{\text{диф}} = W^2/(3D). \qquad (3.52)$$

4) $W \ll L$, высокие частоты. Критерием высокой частоты для диода с тонкой базой оказывается большое значение аргументов гиперболических котангенсов в (3.34), т. е.

$$\left|\frac{W}{\Lambda}\right| = \left|\frac{W}{L}\sqrt{1 + j\omega\tau}\right| \gg 1,$$

или иначе, пренебрегая единицей по сравнению с $\omega\tau$,

$$\frac{W^2}{L^2}\omega\tau \gg 1.$$

Так как аргументы гиперболических котангенсов оказываются большими, расчет параметров диода в данном случае аналогичен расчету при высоких частотах диода с толстой базой. Это относится к тем формулам, в которые еще не введен ток насыщения.

§ 3.7. ГРАФИКИ ЧАСТОТНЫХ ЗАВИСИМОСТЕЙ ПАРАМЕТРОВ ДИОДА

При построении графиков примем одинаковыми плотности тока для диодов с разной толщиной базы. Одинаковыми будем считать и электрофизические параметры материала диодов. Масштаб по осям удобно взять логарифмический (рис. 3.11).

1. **Активное сопротивление** r. Для диода с толстой базой в области низких частот ($\omega\tau \ll 1$) сопротивление не зависит от частоты в соответствии с (3.37). При высоких частотах ($\omega\tau \gg 1$) сопротивление падает обратно пропорционально корню квадратному из частоты (3.43), что в логарифмическом масштабе дает прямую линию с угловым коэффициентом 1/2 (рис. 3.11, *а*).

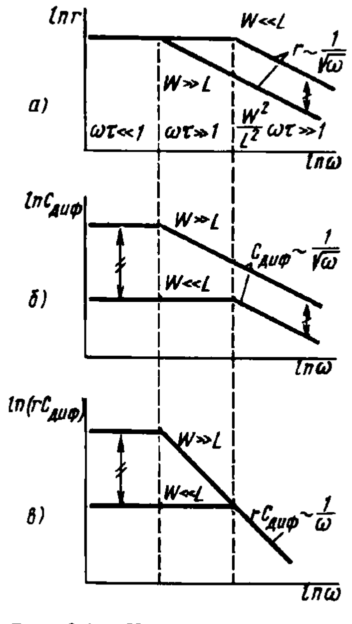


Рис. 3.11. Частотные зависимости активного сопротивления (*а*), диффузионной емкости (*б*) и постоянной времени (*в*) диодов с тонкой и толстой базами

Для диода с тонкой базой низкочастотное значение сопротивления оказывается тем же [ср. (3.49) и (3.37)] и сохраняется постоянным до тех пор, пока не нарушится условие малости аргументов гиперболических котангенсов. При $\frac{W^2}{L^2}\omega\tau \gg 1$ сопротивление падает обратно пропорционально корню квадратному из частоты.

2. **Диффузионная емкость** $C_{диф}$. Диффузионная емкость диода с толстой базой не зависит от частоты при $\omega\tau \ll 1$ (3.39) и обратно пропорциональна корню квадратному из частоты при $\omega\tau \gg 1$ (3.45).

Низкочастотное значение диффузионной емкости диода с тонкой базой значительно меньше, чем диода с толстой базой [ср. (3.51) и (3.39)], так как

$$\frac{W^2}{3D} = \frac{W^2}{3L^2}\tau \ll \frac{\tau}{2} \text{ при } W \ll L.$$

В диапазоне частот, когда $\frac{W^2}{L^2}\omega\tau \gg 1$, диффузионная емкость обратно пропорциональна корню квадратному из частоты.

При высоких частотах постоянные времени любого диода должны быть равны между собой и равны $1/\omega$ в соответствии с (3.46), поэтому на рис. 3.11, *а*, *б* отмечены равные отрезки в диапазоне высоких частот.

3. **Постоянные времени** $rC_{диф}$. Построение графиков для постоянной времени следует из двух предыдущих построений.

При низких частотах постоянная времени диода с толстой базой много больше постоянной времени диода с тонкой базой [ср. (3.40) и (3.52)]. При высоких частотах постоянные времени этих диодов одинаковы (рис. 3.11, *в*).

§ 3.8. ФИЗИЧЕСКИЙ СМЫСЛ ПАРАМЕТРОВ ДИОДА

Низкочастотные значения

1. **Сопротивление диода** *r* представляет собой просто днфференцнальное сопротивлеиие диода, т. е. сопротивленне днода малому переменному току при постоянном смещенин. Чтобы в этом убеднться, надо продифференцнровать выражение ВАХ диода (3.36) н сравиить с (3.37):

$$\frac{dI}{dU} = \frac{q}{kT} I_{\text{нас}} \exp \frac{qU}{kT} = \frac{q}{kT} (I + I_{\text{нас}}) = \frac{1}{r}.$$

2. **Диффузионную емкость** обычно связывают с изменением заряда ннжектированных иосителей при нзменеини напряжения на диоде. Действнтельно, инжектнроваиные иосителн в течение некоторого временн существуют в областях диода, прнмыкающих к *р-п*-переходу. Прн изменеиии напряженяя часть накопленных иеосновных иоснтелей может возвратиться в *р-п*-переход и пройтн через него в соседнюю область. Образующийся прн этом ток аналогичеи емкостному току.

Однако к этому требуются некоторые поясненяя н дополнення. Дело в том, что прн иижекцни области, прнмыкающие к *р-п*-переходу, остаются иейтральными, т. е. инкакой суммариый заряд в ннх ие появляется. Нейтрализация заряда происходнт из-за подхода основиых носителей в те области, куда произошла инжекция неосновных иоснтелей. Нейтрализация устанавливается за очень малый промежуток времеин — порядка временн максвелловской нли диэлектрнческой релаксацин (обычно $10^{-11}...10^{-12}$ с). Так как концентрацня осиовных носнтелей отиосительно велика и необходимое их количество пополияется невыпрямляющим контактом, нейтралнзацня получается практически полной. Следует заметнть, что нейтрализуется ие только заряд в средием по всей областн, но и заряд в каждой точке, т. е. выполняется условне локальной электрнческой иейтральности.

Несмотря на то что прн нижекции прнмыкающне к *р-п*-переходу области ие заряжаются, диффузионную емкость можно связать с зарядом ннжектированиых носителей, так как нижектироваииые неосиовиые иоснтелн н нейтрализующие их основные иосители не нсчезают. Для сравнения вспомним, что и обычиый коидеисатор в целом электрнческн иейтрален. Но в обычиом коиденсаторе положительиый и отрнцательный заряды простраиственно разделеиы (то же самое можио сказать и о *р-п*-переходе прн рассмотренни его барьерной емкостн), в то время как при

инжекции через p-n-переход и положительный, и отрицательный заряды оказываются в одной и той же области и пространственно не разделяются, в результате чего невозможно обнаружить область, где проходят токи смещения.

Следовательно, диффузионную емкость можно связать с изменением заряда инжектированных неосновных носителей, но нельзя связать с прохождением токов смещения. В этом существенное физическое отличие диффузионной емкости от барьерной емкости p-n-перехода и от емкости обычного конденсатора.

Диффузионную емкость можно представить следующим образом:

$$C_{\text{диф}} = |dQ_{\text{инж. эф}}/dU|, \tag{3.53}$$

где $dQ_{\text{инж.эф}}$ — эффективное значение инжектированного заряда.

Здесь абсолютное значение отношения взято для того, чтобы не возникала путаница из-за правила знаков для напряжения, а также из-за того, что инжектированный заряд может быть как положительным, так и отрицательным. Эффективное значение инжектированного заряда следует брать потому, что из-за распределенного характера этого заряда он не весь одинаково участвует в образовании емкости. Поэтому приходится проводить какое-то усреднение.

Чтобы убедиться, что данное определение не противоречит полученным ранее значениям диффузионной емкости, можно найти значение

$$|dQ_{\text{инж}}/dU|,$$

где $Q_{\text{инж}}$ — полный инжектированный заряд.

Полученный при дифференцировании $Q_{\text{инж}}$ по U результат, по-видимому, должен несущественно отличаться от значения диффузионной емкости (3.53). Сравнение $dQ_{\text{инж}}/dU$ с $C_{\text{диф}}$ дает возможность найти и способ выражения эффективного инжектированного заряда через полный заряд инжектированных носителей.

Запишем выражение для полного инжектированного заряда, учитывая, что он складывается в общем случае из заряда дырок, инжектированных в n-область, и электронов, инжектированных в p-область. Тогда в системе координат, принятой на рис. 3.5,

$$Q_{\text{инж}} = qS\left(\int_0^{W_n} \Delta p_n dx + \int_{-W_p}^{0'} \Delta n_p dx \right). \tag{3.54}$$

Применим это соотношение к частным случаям.

а) Диод с толстой базой. Для диода с толстой базой распределение избыточной концентрации дырок в n-области соответствует выражению (3.28). Аналогичное выражение будет и для распределения избыточной концентрации инжектированных в

p-область электронов. При интегрировании положим $W_n = \infty$ и $W_p = \infty$. Тогда

$$Q_{\text{инж}} = qS(p_{n0}L_p + n_{p0}L_n)\left(\exp\frac{qU}{kT} - 1\right)$$

и

$$\left|\frac{dQ_{\text{инж}}}{dU}\right| = \frac{q^2S}{kT}(p_{n0}L_p + n_{p0}L_n)\exp\frac{qU}{kT}. \qquad (3.55)$$

Полученное выражение отличается от выражения для диффузионной емкости диода с толстой базой (3.38) сравнительно небольшим множителем 2, который можно отнести за счет усреднения. Следовательно, для диода с толстой базой $Q_{\text{инж.эф}} = \frac{1}{2}Q_{\text{инж}}$.

б) Диод с тонкой базой. Для диода с тонкой базой распределение избыточной концентрации дырок в n-области соответствует выражению (3.30). Аналогичное выражение будет и для распределения избыточной концентрации инжектированных в p-область электронов. После интегрирования получим

$$Q_{\text{инж}} = \frac{1}{2}\,qS(p_{n0}W_n + n_{p0}W_p)\left(\exp\frac{qU}{kT} - 1\right)$$

и

$$\left|\frac{dQ_{\text{инж}}}{dU}\right| = \frac{q^2S}{2kT}(p_{n0}W_n + n_{p0}W_p)\exp\frac{qU}{kT}. \qquad (3.56)$$

Этот результат всего в 1,5 раза отличается от полученного для диффузионной емкости диода с тонкой базой (3.50), что также можно отнести за счет усреднения.

Приведенные примеры показывают, что диффузионную емкость диода можно связать с изменением заряда инжектированных носителей, но надо учитывать эффективный инжектированный заряд.

Иногда для определения диффузионной емкости используют формулу

$$C_{\text{диф}} = |dQ_{\text{инж}}/dU|.$$

Такой способ вычисления диффузионной емкости неудобен, так как для расчета полной проводимости диода в его эквивалентную схему придется подставлять не полное значение $C_{\text{диф}}$, а некоторую ее часть, зависящую к тому же от конструкции диода (1/2 для диода с толстой базой и 1/1,5 для диода с тонкой базой).

3. Постоянная времени. Если продолжить аналогию полупроводникового диода с конденсатором, то можно выяснить физический смысл постоянной времени $rC_{\text{диф}}$. Для конденсатора постоянная времени показывает, за какое время его заряд уменьшится в e раз, т. е. постоянная времени характеризует время исчезновения заряда конденсатора.

Постоянная времени диода с толстой базой при низкой частоте $rC_{диф}=\tau/2$ (3.40) тоже характеризует время исчезновения заряда. Действительно, τ — время жизни неосновных носителей — как раз и показывает, в течение какого времени концентрация неосновных носителей изменится в e раз из-за рекомбинации.

Для диода с тонкой базой при низкой частоте в соответствии с (3.52) постоянная времени $rC_{диф}=W^2/(3D)$.

Чтобы разобраться в физическом смысле этого результата, вспомним, что в данном случае из-за малой толщины базы процесс рекомбинации носителей в ее объеме уже не является определяющим. Главным теперь оказывается уход неосновных носителей заряда на омический переход с последующей рекомбинацией. Отсюда следует, что время существования заряда в базе теперь будет определяться временем пролета носителей заряда через базу диода.

Чтобы убедиться в этом, выполним расчет *времени пролета неосновных носителей заряда через базу диода*. В общем случае плотность дырочного тока

$$J_p=qpv_p=qp\frac{dx}{dt},$$

отсюда

$$dt=\frac{qp}{J_p}dx.$$

Проинтегрировав последнее выражение в пределах от 0 до $t_{прол}$ по времени и соответственно от 0 до W_n по координате x, получим время пролета:

$$t_{прол}=q\int_0^{W_n}\frac{p}{J_p}dx.$$

Для диода с тонкой базой с учетом (3.30)

$$p_n(x)=p_{n0}+\Delta p_n(x)=p_{n0}+p_{n0}\left(\exp\frac{qU}{kT}-1\right)\left(1-\frac{x}{W_n}\right),$$

а плотность дырочного тока одинаковая во всех сечениях с учетом (3.27) и (3.31)

$$J_p=\frac{qp_{n0}D_p}{W_n}\left(\exp\frac{qU}{kT}-1\right).$$

При расчете для упрощения пренебрежем равновесной концентрацией неосновных носителей, что соответствует большим прямым напряжениям на p-n-переходе. Тогда

$$t_{прол}=q\int_0^{W_n}\frac{W_n-x}{qD_p}dx=\frac{W_n^2}{2D_p}. \qquad (3.57)$$

Сравнение (3.57) с постоянной времени диода с тонкой базой (3.52) подтверждает, что она определяется временем пролета неосновных носителей заряда через базу диода.

Высокочастотные значения

Для высокочастотных значений сопротивления и диффузионной емкости диода характерна их частотная зависимость. Это сильно ограничивает использование таких параметров диода, так как затрудняет расчет частотных характеристик схем с использоваиием полупроводниковых диодов.

Такая частотная зависимость появилась из-за того, что система с распределенными параметрами (диод на высокой частоте) была представлена моделью диода с сосредоточенными параметрами, которая является неудачной для высоких частот. Отсюда следует, что нельзя искать физический смысл высокочастотных параметров полупроводникового диода. Их надо рассматривать как формальные.

§ 3.9. ПРЕДЕЛЫ ПРИМЕНИМОСТИ ЧАСТНЫХ СЛУЧАЕВ РАСЧЕТА ПАРАМЕТРОВ ДИОДА

Предыдущие расчеты были приведены для частных случаев диодов с толстой и тонкой базами, при низкой и высокой частотах. Однако полученные выражения со сравнительно небольшими погрешностями могут применяться практически всегда. Чтобы в этом убедиться, рассмотрим ошибки, которые могут возникать по разным причииам.

1. Размеры базы диода. Частные случаи толщины базы ($W \gg L$ и $W \ll L$) использовались для того, чтобы избавиться от гиперболических функций в формулах. По-видимому, наибольшую погрешность полученные соотношения будут иметь при $W \approx L$.

Пусть $W = L$, тогда

$$\mathrm{cth}(W/L) = \mathrm{cth}\,1 = 1{,}31.$$

Если рассматривать такой диод как диод с толстой базой, то $\mathrm{cth}(W/L) \approx 1$, что дает погрешность 24% от точного значения.

Если же рассматривать такой диод как диод с тонкой базой, то при разложении гиперболического котангенса в ряд $\mathrm{cth}(W/L) \approx \approx (L/W) = 1$, что дает ту же погрешность 24%. Для расчета токов в полупроводниковом диоде эту погрешность можно считать умеренной.

2. Частотные диапазоны. Аналогично сказанному, можно считать, что наибольшая погрешность приближенных расчетов будет при $\omega\tau \approx 1$. Пусть $\omega\tau = 1$, тогда

$$\sqrt{1 + j\omega\tau} = \sqrt{1 + j} = \sqrt[4]{2}\left(\cos\frac{\pi}{8} + j\sin\frac{\pi}{8}\right) = 1{,}1 + j0{,}45.$$

Если принять приближение $\omega\tau \ll 1$, то, раскладывая $\sqrt{1+j}$ в биномиальный ряд, получим $\sqrt{1+j} \approx 1{,}0 + j0{,}5$, что дает погрешность около 10% для действительной и мнимой частей.

Если же использовать методику расчета параметров такого диода, соответствующую высокой частоте, то

$$\sqrt{1+j} \approx \sqrt{j} = \sqrt{\cos 90° + j \sin 90°} =$$
$$= \cos 45° + j \sin 45° = 0{,}71 + j0{,}71,$$

что дает погрешность действительиой части около 35%, а мнимой — около 60%, т. е. получается правильный порядок.

Таким образом, судя по максимально возможным погрешностям, для практических расчетов можно ограничиться только рассмотренными частными случаями.

§ 3.10. ГЕНЕРАЦИЯ И РЕКОМБИНАЦИЯ НОСИТЕЛЕЙ ЗАРЯДА В ЭЛЕКТРОННО-ДЫРОЧНОМ ПЕРЕХОДЕ

Толщина *p-n*-перехода в диодах значительно меньше диффузионной длины носителей заряда. Однако процессы генерации и рекомбинации носителей в *p-n*-переходе идут более интенсивно, чем в прилегающих к нему областях. Поэтому составляющие тока полупроводникового диода, связанные с генерацией и рекомбинацией носителей в *p-n*-переходе, могут быть существенными.

При обратном напряженни на диоде образующиеся из-за тепловой генерации носители разного знака растаскиваются электрическим полем перехода в разные стороны и уже не могут рекомбинировать (рис. 3.12, *а*). Именно поэтому процесс генерации носителей в переходе не уравновешивается процессом рекомбинации. Обратный ток, вызванный генерацией носителей в *p-n*-переходе, называют *генерационным током* ($I_{ген}$). Следует, однако, помнить, что ток насыщения, связанный с экстракцией неосновных носителей заряда и рассмотренный в § 3.2 и 3.4, также вызван генерацией неосновных носителей заряда, но генерацией в прилегающих к переходу областях.

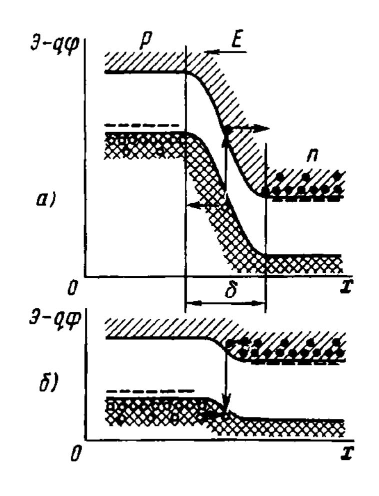


Рис. 3.12. Генерация (*а*) и рекомбинация (*б*) носителей заряда в *p-n*-переходе

Чтобы выяснить влияние генерационного тока на обратный ток через днод, проведем качественное сравнение плотности тока насыщения и плотности тока генерации. Предположим, что диод имеет снмметричный *p-n*-переход и все электрофизические параметры прнлегающих к переходу областей также одинаковы,

т. е. $p_{n0}=n_{p0}$, $L_p=L_n=L$, $\tau_p=\tau_n=\tau$. Тогда в соответствии с (3.1) в диоде с $W\gg L$ плотность тока насыщения

$$J_{\text{нас}}=2qp_{n0}L/\tau.$$

На основании физики процессов образования тока насыщения и тока генерации можно по аналогии с предыдущим выражением записать соотношение, показывающее, чему пропорционален ток генерации. При этом следует учесть, что если в образовании тока насыщения принимали участие неосновные носители заряда с концентрацией p_{n0}, возникающие в объеме толщиной $2L$, то в образовании тока генерации принимают участие носители с концентрацией n_i, возникающие в p-n-переходе толщиной δ. Поэтому

$$J_{\text{ген}}\sim qn_i\delta/\tau.$$

Тогда отношение плотности тока генерации к плотности тока насыщения с учетом (1.19) пропорционально следующим величинам:

$$\frac{J_{\text{ген}}}{J_{\text{нас}}}\sim\frac{n_i\delta}{p_{n0}L}=\frac{n_{n0}\delta}{n_iL}.$$

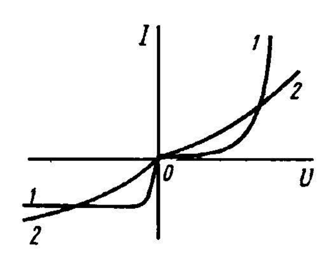


Рис. 3.13. Зависимости от напряжения составляющих тока через диод, вызванных только процессами инжекции и экстракции носителей заряда через p-n-переход (*1*) и процессами рекомбинации и генерации носителей заряда в p-n-переходе (*2*)

Таким образом, во-первых, процесс генерации носителей в p-n-переходе следует учитывать для диодов, изготовленных из полупроводника с большой шириной запрещенной зоны, т. е. с малой собственной концентрацией носителей n_i. Чтобы лучше осознать этот вывод, надо учесть, что плотность тока насыщения в соответствии с (3.2) пропорциональна n_i^2. Так, в кремниевых диодах обратный ток определяется током генерации, а в германиевых диодах — током насыщения.

Во-вторых, с увеличением обратного напряжения по абсолютному значению плотность тока генерации также растет в связи с увеличением толщины p-n-перехода δ. Поэтому в диодах с преобладающим током генерации нет участка постоянного тока на обратной ветви ВАХ (рис. 3.13).

В-третьих, соотношение между током генерации и током насыщения изменяется при изменении температуры — с повышением температуры оба тока растут, но ток насыщения растет быстрее (как n_i^2), а ток генерации — медленнее (как n_i). Значит, с повышением температуры относительная роль тока генерации уменьшается.

В-четвертых, отношение $J_{\text{ген}}/J_{\text{нас}}$ изменяется при изменении концентрации примесей ($n_{n0}\approx N_{\text{дп}}$) — с увеличением концентрации примеси в базе диода относительная роль тока генерации также увеличивается.

При прямом напряжении на диоде носители разных знаков подходят к p-n-переходу. Если прямое напряжение мало́, то высота потенциального барьера на переходе велика и основная часть носителей не может преодолеть потенциальный барьер перехода, но вблизи середины p-n-перехода может происходить их рекомбинация (рис. 3.12, *б*). Составляющую прямого тока, связанную с процессом рекомбинации носителей в p-n-переходе, называют *рекомбинационным током*. Надо понимать условность этого термина, так как прямой ток, связанный с инжекцией неосновных носителей в прилегающие к переходу области (см. § 3.2), также сопровождается рекомбинацией инжектированных носителей либо в базе диода, либо на омическом переходе диода. При больших для диода прямых напряжениях высота потенциального барьера на переходе небольшая. Поэтому прямой ток при больших прямых напряжениях будет вызван в основном инжекцией носителей через уменьшенный потенциальный барьер перехода. Следовательно, рекомбинационный ток может сказываться на значении прямого тока через диод только при малых прямых напряжениях (рис. 3.13).

Плотность рекомбинационного тока в диоде с симметричным p-n-переходом

$$J_{\text{рек}}=\frac{qn_i\delta}{\tau}\,\frac{kT}{q(\varphi_{\text{кон}}-U)}\left[\exp\frac{qU}{2kT}-\exp\left(-\frac{qU}{2kT}\right)\right]. \qquad (3.58)$$

Тогда для диода с $W \gg L$, учитывая (3.2) и (3.27), получим

$$\frac{J_{\text{рек}}}{J_{\text{инж}}}\approx\frac{N\delta}{2n_iL}\,\frac{kT}{q(\varphi_{\text{кон}}-U)}\exp\left(-\frac{qU}{2kT}\right).$$

Отношение значений рекомбинационного тока к инжекционному зависит от ширины запрещенной зоны полупроводника. В диодах из материала с большой шириной запрещенной зоны (кремниевый диод) из-за большой высоты потенциального барьера инжекция затруднена. Поэтому прямой ток при малых прямых напряжениях будет определяться рекомбинационным током.

§ 3.11. ЛАВИННЫЙ ПРОБОЙ

Механизм лавинного пробоя и основные определения

Обратное напряжение, приложенное к диоду, обычно падает на выпрямляющем электрическом переходе диода. При больших для конкретного диода обратных напряжениях происходит пробой выпрямляющего электрического перехода. *Пробой выпрямляющего электрического перехода* (и соответственно *пробой диода*) — это явление резкого увеличения дифференциальной проводимости выпрямляющего перехода при достижении обратным напряжением критического для данного прибора значения. В зависимости от физических явлений, приводящих к пробою, различают лавинный, туннельный и тепловой пробои.

Лавинный пробой выпрямляющего электрического перехода — это пробой, вызванный лавинным размножением носителей заряда под действием сильного электрического поля. Лавинное размножение носителей заряда происходит в результате того, что они, проходя через выпрямляющий переход при обратном напряжении, приобретают в сильном электрическом поле на длине свободного пробега дополнительную энергию, достаточную для образования новых электронно-дырочных пар носителей заряда посредством ударной ионизации атомов полупроводника.

Процесс ударной ионизации характеризуют коэффициентами ударной ионизации α_n и α_p (см. § 1.10), которые в сильной степени зависят от напряженности электрического поля. Поэтому коэффициенты ударной ионизации для электронов и дырок обычно считают равными.

Чтобы количественно охарактеризовать увеличение тока из-за процесса ударной ионизации в выпрямляющем переходе, вводят *коэффициенты лавинного размножения* M_n и M_p, показывающие, во сколько раз увеличивается ток данных носителей в результате ударной ионизации. Другими словами, коэффициент лавинного размножения — это отношение тока данных носителей заряда (например, электронов), выходящих из выпрямляющего перехода, к току тех же носителей, входящих в переход. В связи с предположением о равенстве коэффициентов ударной ионизации автоматически получается равенство и коэффициентов лавинного размножения: $M_n = M_p = M$.

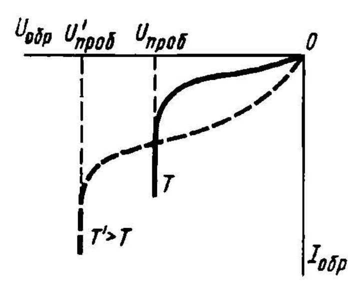


Рис. 3.14. ВАХ диода при лавинном пробое для разных температур

Наконец, параметром диода, характеризующим явление пробоя его выпрямляющего перехода, является *пробивное напряжение диода* — напряжение, при котором происходит неограниченное возрастание тока (рис. 3.14). Формально при пробивном напряжении $M \to \infty$. В производственных условиях пробивное напряжение диода определяют значением обратного напряжения, вызывающего пробой выпрямляющего перехода, при котором обратный ток достигает заданного значения.

Связь коэффициента лавинного размножения с коэффициентом ударной ионизации и вывод условия лавинного пробоя

Обычно в качестве выпрямляющего перехода в различных диодах используют *p-n*-переход. Поэтому дальнейшее рассмотрение лавинного пробоя проведем для *p-n*-перехода. Для расчета связи

параметра, характеризующего физический процесс, — коэффициеита удариой ионизации — с параметром, характеризующим p-n-переход при удариой иоиизации, — коэффициеитом лавииного размиожеиия — воспользуемся уравиеиием иепрерывиости, иапример для электроиов, которое имеет вид, аиалогичиый виду уравнеиия иепрерывиости для дырок (3.5):

$$\frac{\partial n}{\partial t}=\frac{1}{q}\operatorname{div}J_n-R_n+G_n. \qquad (3.59)$$

Если при рассмотрении установившегося процесса ударной ионизации ($\partial n/\partial t=0$) пренебречь рекомбииацией в p-n-переходе при большом обратиом иапряжеиии иа диоде ($R_n=0$), то уравиение иепрерывности для одиомериой модели диода примет вид

$$-\frac{1}{q}\frac{dJ_n}{dx}=G_n. \qquad (3.60)$$

В даином случае скорость иетепловой генерации электронов (G_n) учитывает их геиерацию под действием сильиого электрического поля. Скорость иетепловой генерацни численио равна числу иосителей, образующнхся в едииицу времеии в единице объема. Чтобы выразить скорость нетепловой генерацин через коэффициеиты ударной ионизации, рассмотрим единичный объем, имеющий длину в направлении прохождения иосителей заряда и площадь попёречиого сечения, равные едиииде. Тогда каждый проходящий через этот объем носитель образует в едииичном объеме количество иосителей, равное коэффициеиту ионизации α. В единицу времеии через рассматриваемый объем проходнт число носителей, равное J/q. Следовательио, если учесть геиерацию электроиов (или дырок) в результате иоиизацни атомов электронами и дырками, получим

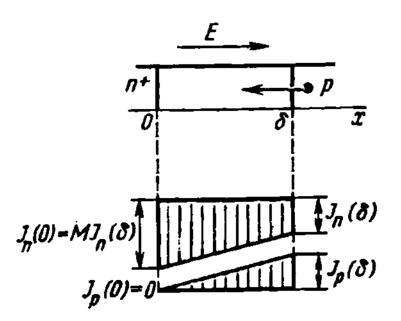


Рис. 3.15. Плотиости электронного и дырочиого токов при лавинном размножении в несимметричном n^+-p-переходе

$$G_n=G_p=\alpha\frac{J_n}{q}+\alpha\frac{J_p}{q}. \qquad (3.61)$$

Подставим (3.61) в (3.60). Тогда.

$$-\frac{dJ_n}{dx}=\alpha(J_n+J_p)=\alpha J. \qquad (3.62)$$

Для решения дифференциального уравнеиия (3.62) выясним граиичные условия. Для иесимметричиого p-n-перехода (например, n^+-p) можно счнтать, что в переход входит только поток

электронов, соответствующий плотности электронного тока $J_n(\delta)$ (рис. 3.15). В соответствии с определением коэффициента лавинного размножения выходящая из перехода плотность электронного тока равна $MJ_n(\delta)$. Эта же плотность тока является полной плотностью тока через рассматриваемый переход, так как через левую границу несимметричного n^+-p-перехода дырочная составляющая тока не проходит. Таким образом, граничные условия принимают вид

$$\begin{aligned} &\text{при } x=0 \quad J_n(0)=MJ_n(\delta)=J;\; J_p=J_p(0)=0;\\ &\text{при } x=\delta \quad J_n=J_n(\delta);\; J_p=J_p(\delta). \end{aligned} \tag{3.63}$$

С учетом проведенных рассуждений уравнение (3.62) примет вид

$$-\frac{dJ_n}{dx}=\alpha MJ_n(\delta),$$

или

$$-dJ_n=\alpha MJ_n(\delta)dx. \tag{3.64}$$

При интегрировании дифференциального уравнения (3.64) надо согласовать пределы интегрирования, используя граничные условия (3.63), которые наглядно представлены на рис. 3.15. Тогда

$$-\int\limits_{J_n(\delta)}^{MJ_n(\delta)} dJ_n = MJ_n(\delta)\int\limits_0^0 \alpha\, dx$$

или

$$-MJ_n(\delta)+J_n(\delta)=MJ_n(\delta)\int\limits_\delta^0 \alpha\, dx.$$

В результате

$$1-\frac{1}{M}=\int\limits_0^\delta \alpha\, dx. \tag{3.65}$$

Соотношение (3.65) отражает связь коэффициента лавинного размножения с коэффициентом ударной ионизации. При пробое p-n-перехода коэффициент лавинного размножения $M\to\infty$. Тогда

$$\int\limits_0^\delta \alpha\, dx=1. \tag{3.66}$$

Уравнение (3.66) является *условием лавинного пробоя p-n-перехода.*

Расчет коэффициента лавинного размножения и пробивного напряжении при лавинном пробое резкого электронно-дырочного перехода

Распределение напряженности электрического поля в резком несимметричном n^+-p-переходе линейно (2.18), и практически вся область объемного заряда расположена в слаболегированной области полупроводника (2.22). Тогда в соответствии с (2.18) и с учетом изменения принятого начала координат ($\delta \approx \delta_p$)

$$|E| = \frac{qN_{ap}}{\varepsilon\varepsilon_0}(\delta - x).$$

Учитывая соотношение между коэффициентами лавинного размножения и ударной ионизации (3.65) и принимая аппроксимацию зависимости коэффициента ударной ионизации от напряженности электрического поля (1.33), получим

$$1 - \frac{1}{M} = \int_0^\delta \alpha dx = A\left(\frac{qN_{ap}}{\varepsilon\varepsilon_0}\right)^m \int_0^\delta (\delta - x)^m dx =$$

$$= A\left(\frac{qN_{ap}}{\varepsilon\varepsilon_0}\right)^m \frac{\delta^{m+1}}{m+1}. \qquad (3.67)$$

Подставим в последнее выражение соотношение для толщины резкого несимметричного p-n-перехода (2.27), в котором можно пренебречь значением контактной разности потенциалов по сравнению с обратным напряжением ($\varphi_{кон} \ll |U|$). Тогда

$$1 - \frac{1}{M} = \frac{A}{m+1}\left(\frac{qN_{ap}}{\varepsilon\varepsilon_0}\right)^m \left(\frac{2\varepsilon\varepsilon_0}{qN_{ap}}\right)^{\frac{m+1}{2}} |U|^{\frac{m+1}{2}} \qquad (3.68)$$

С помощью этого соотношения можно рассчитать коэффициент лавинного размножения, зная концентрацию примесей в базе диода и другие параметры материала базы.

При лавинном пробое (при $U_{проб}$) коэффициент лавинного размножения $M \to \infty$. Поэтому

$$1 = \frac{A}{m+1}\left(\frac{qN_{ap}}{\varepsilon\varepsilon_0}\right)^m \left(\frac{2\varepsilon\varepsilon_0}{qN_{ap}}\right)^{\frac{m+1}{2}} U_{проб}^{\frac{m+1}{2}} \qquad (3.69)$$

Разделив (3.68) на (3.69), получим

$$1 - 1/M = (U/U_{проб})^{\frac{m+1}{2}}$$

т. е.

$$M = \frac{1}{1 - (U/U_{\text{проб}})^{\frac{m+1}{2}}},$$

или

$$M = \frac{1}{1 - (U/U_{\text{проб}})^b}, \qquad (3.70)$$

где $b = (m + 1)/2$ — коэффициент, различный для диодов из разиых материалов ($b = 2...6$).
При $U \to 0$ $M \to 1$, при $U \to U_{\text{проб}}$ $M \to \infty$ (рис. 3.16).

Из выражеиия (3.69) пробивиое напряжение при лавинном пробое

$$U_{\text{проб}}^{\frac{m+1}{2}} = \frac{m+1}{A}\left(\frac{1}{2}\right)^{\frac{m+1}{2}}\left(\frac{\varepsilon\varepsilon_0}{q}\right)^{m-\frac{m+1}{2}}\left(\frac{1}{N_{ap}}\right)^{m-\frac{m+1}{2}},$$

или

$$U_{\text{проб}} = \left[\frac{1}{2}\left(\frac{m+1}{A}\right)^{\frac{2}{m+1}}\left(\frac{\varepsilon\varepsilon_0}{q}\right)^{\frac{m-1}{m+1}}\right]\frac{1}{N_{ap}^{\frac{m-1}{m+1}}} \qquad (3.71)$$

Здесь выражение в квадратных скобках включает в себя только величииы, которые можио считать постоянными для даиного материала. Тогда, вводя новые обозначения, получим

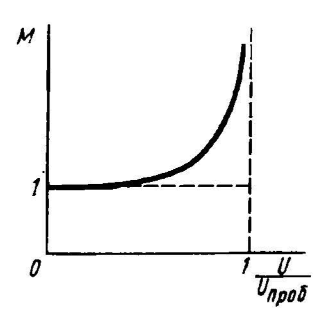


Рис. 3.16.Зависимость коэффициента лавинного размножения от напряжения на p-n-переходе

$$U_{\text{проб}} = B_1\frac{1}{N^k}, \qquad (3.72)$$

где N — концентрация примесей в слаболегированной области, т. е. в базе диода; $k = (m - 1)/(m + 1)$.

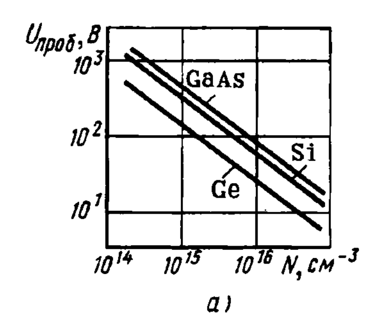
Экспериметы подтверждают соотношение (3.72). В системе координат $\lg U_{\text{проб}}$ от $\lg N$ экспериментальиые точки ложатся на прямую линию (рис. 3.17, a). При этом для несимметричных p-n-переходов p^+-n и n^+-p, сформированных в одном и том же полупроводнике, зависимости $\lg U_{\text{проб}} = f(\lg N)$ совпадают. Этот экспериментальный факт доказывает правомерность допущения равенства коэффициентов ударной ионизации для электронов и дырок, что было принято в начале этого параграфа.

Однако чаще бывает известна не концентрация примесей в базе диода, а ее удельиое сопротивление, т. е. удельиое сопро-

тивление исходного полупроводника. Поэтому, учитывая связь концентрации примеси или концентрации основных носителей с удельным сопротивлением, соотношение (3.72) удобнее записать в виде

$$U_{\text{проб}} = B\varrho^{a}. \qquad (3.73)$$

Для кремниевых p^+-n-переходов $U_{\text{проб}} = 96\varrho^{0,78}$; n^+-p-переходов $U_{\text{проб}} = 48\varrho^{0,78}$, для германиевых p^+-n-переходов $U_{\text{проб}} = 100\varrho^{0,8}$; n^+-p-переходов $U_{\text{проб}} = 55\varrho^{0,8}$, где ϱ — удельное сопротивление базы, Ом·см.



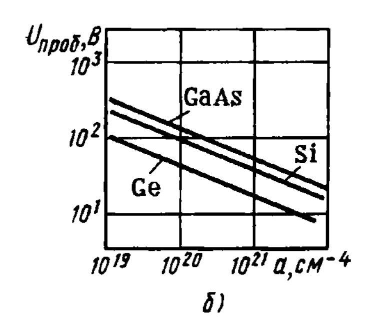


Рис. 3.17. Зависимости пробивного напряжения при лавинном пробое от концентрации примесей в базе диода с несимметричным резким p-n-переходом (а) и от градиента концентрации примесей в плавном p-n-переходе (б)

Эмпирические коэффициенты B и a различны не только для диодов из разных материалов, но и для диодов из одного и того же материала с разными резкими p-n-переходами (p^+-n и n^+-p). Связано это различие в коэффициентах с тем, что подвижность электронов отличается от подвижности дырок в одном и том же материале.

Таким образом, пробивное напряжение резких несимметричных p-n-переходов определяется концентрацией примесей в слаболегированной области или ее удельным сопротивлением, так как от этих величин зависит толщина p-n-перехода.

Толщина плавных p-n-переходов зависит от градиента концентрации примесей (2.33). Поэтому пробивное напряжение плавных p-n-переходов определяется градиентом концентрации примесей (рис. 3.17, *б*).

С повышением температуры уменьшается длина свободного пробега носителей заряда, а значит, и энергия, которую носитель заряда может приобрести на длине свободного пробега в электрическом поле. Следовательно, повышение температуры приводит к увеличению пробивного напряжения при лавинном пробое (см. рис. 3.14).

Особенности лавинного пробоя в электронно-дырочных переходах с дефектами

Электронно-дырочные переходы реальных диодов часто имеют различные дефекты: нарушения кристаллической решетки, инородные включения и т. п. Такие неоднородности, независимо от их характера, приводят к нарушениям картины электрического поля, к появлению областей с повышенной напряженностью. В тех местах, где напряженность электрического поля оказывается больше, развивается пробой. После развития пробоя в районе дефекта образуется шнур с повышенной удельной проводимостью. Однако из-за малого сечения шнура последовательно с ним оказывается включенным большое так называемое сопротивление растекания, сосредоточенное в примыкающих к переходу областях (рис. 3.18). Поэтому ток при лавинном пробое по дефектам ограничен большим сопротивлением растекания.

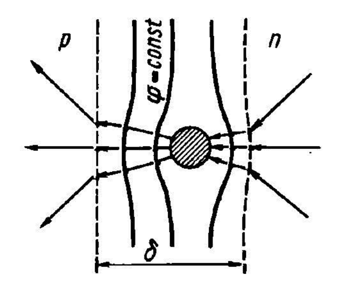


Рис. 3.18. Искажение линий тока и эквипотенциалей в p-n-переходе и в прилегающих к нему областях, связанные с проводящим дефектом в p-n-переходе

Рис. 3.19. ВАХ диода при пробое по дефектам

В связи с тем что количество различных дефектов в p-n-переходе может быть большим и они по своим свойствам как-то распределены статистически, пробой по каждому из них приводит к некоторому росту тока через диод. Таким образом, при лавинном пробое p-n-перехода по дефектам получается «мягкая» обратная ветвь ВАХ (рис. 3.19).

С явлениями лавинного пробоя по дефектам связано также, по-видимому, то, что экспериментальные данные пробивных напряжений и коэффициентов лавинного размножения получаются с большим разбросом.

§ 3.12. ТУННЕЛЬНЫЙ ПРОБОЙ

Туннельным пробоем p-n-перехода называют электрический пробой перехода, вызванный квантово-механическим туннелированием носителей заряда сквозь запрещенную зону полупроводника без изменения их энергии (см. рис. 1.12). Туннелирование электронов возможно при условии, если ширина потенциального барьера Δ, который необходимо преодолеть электронам, достаточно мала. При одной и той же ширине запрещенной зоны (для одного и того же материала) ширина потенциального барьера определяется напряженностью электрического поля, т. е. наклоном энергетических уровней и зон (см. рис. 1.2). Следовательно, условия для туннелирования возникают только при определенной напряженности электрического поля или при определенном напряжении на p-n-переходе — при пробивном напряжении. Значение этой критической напряженности электрического поля составляет примерно $8 \cdot 10^5$ В/см для кремниевых переходов и $3 \cdot 10^5$ В/см для германиевых. Так как вероятность туннелирования очень сильно зависит от напряженности электрического поля, то внешне туннельный эффект проявляется как пробой диода (рис. 3.20).

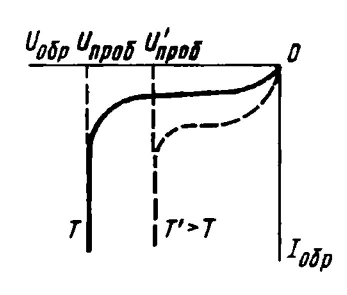


Рис. 3.20. ВАХ диода при туннельном пробое для разных температур

Рассмотрим в качестве примера расчет пробивного напряжения при туннельном пробое резкого близкого к симметричному p-n-перехода.

Наибольшее значение напряженности электрического поля в таком переходе существует на металлургическом контакте [см. § 2.6 и, в частности, формулу (2.18)], т. е. при $x = 0$

$$|E|_{\max} = \frac{qN_{ap}}{\varepsilon\varepsilon_0}\delta_p.$$

Воспользовавшись уравнением (2.23), получим

$$|E|_{\max} = \frac{qN_{ap}}{\varepsilon\varepsilon_0}\frac{N_{дn}}{N_{ap}+N_{дn}}\delta.$$

С учетом соотношения (2.26) для толщины резкого p-n-перехода получим

$$E^2_{\max} = \frac{2q}{\varepsilon\varepsilon_0}\frac{N_{ap}N_{дn}}{N_{ap}+N_{дn}}(\varphi_{кон} - U).$$

При $E_{\max} = E_{кр}$ напряжение на p-n-переходе будет пробивным. Поэтому

$$U_{проб} = \frac{\varepsilon\varepsilon_0 E^2_{кр}}{2q}\frac{N_{ap}+N_{дn}}{N_{ap}N_{дn}} - \varphi_{кон},$$

или

$$U_{проб} = \frac{\varepsilon\varepsilon_0 E^2_{кр}}{2q}\left(\frac{1}{N_{ap}} + \frac{1}{N_{дn}}\right) - \varphi_{кон}. \qquad (3.74)$$

Следовательно, при туннельном пробое пробивное напряжение оказывается обратно пропорциональным концентрации примесей в первой степени.

Туннельный пробой может происходить только в *p-n*-переходах, изготовленных в полупроводниках с большой концентрацией примесей, так как для туннелирования необходимы малая толщина потенциального барьера и, следовательно, малая ширина перехода. При малой ширине перехода пробивные напряжения его оказываются также малыми. Поэтому пренебрегать контактной разностью потенциалов в выражении (3.74) не следует, так как она может быть сравнима с пробивным напряжением.

Если в (3.74) от концентраций примесей перейти к удельным сопротивлениям и все-таки пренебречь контактной разностью потенциалов, то получим широко известное соотношение

$$U_{\text{проб}} = C_p \varrho_p + C_n \varrho_n. \tag{3.75}$$

Для кремниевых *p-n*-переходов после подстановки числовых значений параметров пробивное напряжение при туннельном пробое

$$U_{\text{проб}} = 200\varrho_n + 73\varrho_p;$$

для германиевых переходов

$$U_{\text{проб}} = 190\varrho_n + 94\varrho_p,$$

где ϱ_n и ϱ_p — удельные сопротивления *n*- и *p*-областей, прилегающих к переходу, Ом·см.

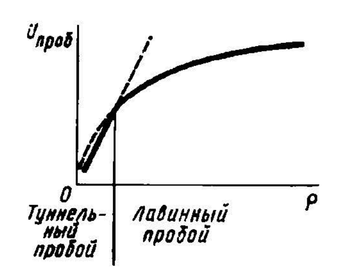


Рис. 3.21. Зависимость пробивного напряжения при туннельном и лавинном пробое от удельного сопротивления базы диода

С повышением температуры у большинства полупроводников ширина запрещенной зоны уменьшается. Следовательно, при этом уменьшается и толщина потенциального барьера при той же напряженности электрического поля, что приводит к увеличению вероятности туннелирования носителей сквозь потенциальный барьер. Поэтому пробивное напряжение при туннельном пробое уменьшается с увеличением температуры (рис. 3.20).

Так как для туннельного пробоя необходима малая толщина *p-n*-перехода, он наблюдается в диодах, изготовленных на основе полупроводников с большой концентрацией примесей. Напряжения туннельного пробоя поэтому не превосходят нескольких вольт. В диодах из полупроводников с меньшей концентрацией примесей и, значит, с большей толщиной *p-n*-перехода пробой будет происходить при больших напряжениях и будет иметь лавинный характер (рис. 3.21).

§ 3.13. ТЕПЛОВОЙ ПРОБОЙ

Вольт-амперная характеристика с учетом тепловыделения в диоде

Тепловой пробой диода — это пробой, развитие которого обусловлено выделением в выпрямляющем электрическом переходе теплоты вследствие прохождения тока через переход. При подаче на диод обратного напряжения практически все оно падает на *p-n*-переходе, через который идет, хотя и небольшой, обратный ток. Выделяющаяся на *p-n*-переходе мощность

$$P_{выд} = U_{обр} I_{обр} \qquad (3.76)$$

вызывает разогрев *p-n*-перехода и прилегающих к нему областей полупроводника.

Отводимая от *p-n*-перехода мощность в результате теплопроводности и дальнейшего рассеяния теплоты в окружающую среду пропорциональна перегреву *p-n*-перехода ($T - T_{окр}$) и обратио пропорциональна тепловому сопротивлению коиструкции диода $R_{T\,пс}$:

$$P_{отв} = \frac{T - T_{окр}}{R_{T\,пс}}. \qquad (3.77)$$

Таким образом, тепловое сопротивление диода между *p-n*-переходом и окружающей средой определяется мощностью, отводимой от перехода в окружающую среду при разиости температур между ними в один кельвин.

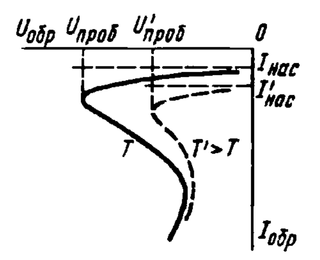


Рис. 3.22. ВАХ диода с учетом выделения теплоты в *p-n*-переходе при разных температурах окружающей среды

Через некоторое время после подачи иа диод обратного напряжеиия устанавливается тепловое равновесие между выделяемой и отводимой теплотой. При этом $P_{выд} = P_{отв}$, или

$$U_{обр} I_{обр} = \frac{T - T_{окр}}{R_{T\,пс}}. \qquad (3.78)$$

Учтем теперь зависимость обратиого тока от температуры. Для простейшего случая, когда обратный ток определяется экстракцией неосиовных иосителей в диоде с толстой базой [см. (3.2) или (3.29)],

$$I_{обр} = J_{нас} S = q S n_i^2 \left(\frac{L_p}{N_{д n} \tau_p} + \frac{L_n}{N_{а p} \tau_n} \right).$$

С учетом (1.15), пренебрегая слабой зависимостью от температуры концентрации примесей, диффузиоииой длины и времеии жизни, получим

$$I_{обр} = I_{\infty} \exp\left(- \frac{\Delta Э}{kT} \right), \qquad (3.79)$$

где I_∞ имеет формальный смысл тока при бесконечно большой температуре.

Подставим (3.79) в уравнение теплового баланса (3.78). Тогда

$$U_{обр} = \frac{T - T_{окр}}{R_{T\,пс} I_\infty} \exp \frac{\Delta Э}{kT}. \tag{3.80}$$

Полученное соотношение показывает связь обратного напряжения на диоде с температурой p-n-перехода. Соотношения (3.79) и (3.80) представляют собой систему двух уравнений обратной ветви ВАХ диода, выраженной в параметрическом виде, через параметр T, т. е. с учетом тепловыделения в p-n-переходе. Общий вид такой кривой показан на рис. 3.22. Кривая идет не из начала координат, так как обратный ток был принят равным току насыщения, т. е. нас не интересовала начальная часть характеристики.

На кривой могут быть два экстремума: максимум и минимум напряжения. Между этими экстремумами находится участок ВАХ с отрицательным дифференциальным сопротивлением. Возможность существования отрицательного дифференциального сопротивления объясняется тем, что с увеличением выделяемой мощности растет температура, а следовательно, увеличиваются ток и выделяемая мощность. Таким образом, в диоде возникает внутренняя положительная обратная связь, которая и может привести к появлению теплового пробоя и отрицательного дифференциального сопротивления.

Проведем исследование экстремумов ВАХ. Для этого надо продифференцировать уравнение (3.80) по температуре и приравнять полученное выражение нулю:

$$\frac{dU_{обр}}{dT} = \frac{1}{R_{T\,пс} T_\infty} \left[\exp \frac{\Delta Э}{kT} + \right.$$

$$\left. + (T - T_{окр}) \left(\exp \frac{\Delta Э}{kT} \right) \frac{\Delta Э}{k} \left(- \frac{1}{T^2} \right) \right] = 0,$$

или

$$T^2 - \frac{\Delta Э}{k} T + \frac{\Delta Э}{k} T_{окр} = 0.$$

Отсюда температура p-n-перехода диода, соответствующая экстремальным напряжениям на диоде,

$$T = \frac{\Delta Э}{2k} \pm \sqrt{\left(\frac{\Delta Э}{2k} \right)^2 - \frac{\Delta Э}{k} T_{окр}}. \tag{3.81}$$

Диод будет обладать отрицательным дифференциальным сопротивлением при обратном включении, а обратная ветвь ВАХ будет иметь два экстремальных напряжения при условии, если корни последнего уравнения будут действительными и разными, т. е. если

$$\left(\frac{\Delta Э}{2k} \right)^2 > \frac{\Delta Э}{k} T_{окр},$$

или

$$\Delta Э > 4kT_{окр}. \qquad (3.82)$$

Последнее неравенство справедливо для всех полупроводников, применяемых для изготовления полупроводниковых диодов.

Расчет пробивного напряжения при тепловом пробое

Чтобы оценить влияние разогрева на обратный ток диода, введем понятие температурного коэффициента обратного тока (по аналогии с другими температурными коэффициентами — ТКρ, ТКR, ТКε и т. д.):

$$\alpha = \text{ТК}I = \frac{1}{I}\frac{dI}{dT}. \qquad (3.83)$$

Подставим в (3.83) аппроксимацию обратного тока (3.79). Тогда

$$\alpha = \frac{1}{I_{\infty}\exp\left(-\frac{\Delta Э}{kT}\right)} \frac{I_{\infty}\left[\exp\left(-\frac{\Delta Э}{kT}\right)\right]\Delta Э}{kT^2} = \frac{\Delta Э}{kT^2}. \qquad (3.84)$$

В дифференциальном уравнении (3.83) можно разделить переменные, а потом провести интегрирование, согласовав пределы интегрирования:

$$\int_{I_{окр}}^{I} \frac{dI}{I} = \alpha \int_{T_{окр}}^{T} dT,$$

где $I_{окр}$ — обратный ток через диод при начальной температуре $T_{окр}$.

При интегрировании пренебрежем температурной зависимостью α, вынося этот параметр за знак интеграла. В этом случае окончательный результат можно будет считать справедливым в первом приближении только при небольших перегревах p-n-перехода относительно среды, окружающей диод. Тогда в результате интегрирования получим

$$\alpha(T - T_{окр}) = \ln\frac{I}{I_{окр}},$$

или

$$\exp\left[\alpha(T - T_{окр})\right] = \frac{I}{I_{окр}},$$

т. е.

$$I = I_{окр}\exp\left[\alpha(T - T_{окр})\right]. \qquad (3.85)$$

Если теперь подставить выражение обратного тока через температурный коэффициент этого тока (3.85) в уравнение теплового баланса (3.78), то

$$U = \frac{T - T_{окр}}{R_{T\,пс} I_{окр} \exp[\alpha(T - T_{окр})]}. \quad (3.86)$$

При тепловом пробое $dU/dT = 0$. Поэтому после дифференцирования формулы (3.86) и сокращений получим

$$1 - \alpha(T - T_{окр}) = 0.$$

Отсюда температура *p-n*-перехода при тепловом пробое

$$T = T_{окр} + 1/\alpha. \quad (3.87)$$

Для полупроводниковых диодов значение температурного коэффициента обратного тока обычно около 0,1 K^{-1}, т. е. при тепловом пробое температура *p-n*-перехода превышает температуру окружающей среды всего примерно на 10 К. Именно из-за малого перегрева *p-n*-перехода в начале развития теплового пробоя можно считать температурный коэффициент обратного тока величиной неизменной при изменении температуры и вынести его за знак интеграла. Конечно, при развитии теплового пробоя с дальнейшим увеличением обратного тока температура *p-n*-перехода может значительно повыситься — вплоть до плавления полупроводникового материала.

Воспользовавшись полученым значением температуры *p-n*-перехода при тепловом пробое (3.87) и подставив его в (3.86), получим

$$U_{проб} = \frac{1}{\alpha R_{T\,пс} I_{окр} e} \quad (3.88)$$

Следовательно, пробивное напряжение при тепловом пробое диода определяется его обратным током, температурным коэффициентом обратного тока и тепловым сопротивлением. Особое внимание следует обратить на сильную зависимость напряжения теплового пробоя от температуры окружающей среды. С увеличением температуры окружающей среды пробивное напряжение при тепловом пробое в соответствии с (3.88) и (3.79) уменьшается (рис. 3.22). Пробивное напряжение уменьшается, во-первых, в связи с увеличением выделяющейся мощности при тех же обратных напряжениях и, во-вторых, из-за ухудшения теплоотвода от *p-n*-перехода.

Так как пробивное напряжение при тепловом пробое зависит от обратного тока через диод при температуре окружающей среды, то в диодах с большими обратными токами даже при комнатных температурах создаются условия для теплового пробоя и он наступает раньше, чем лавинный пробой. Это справедливо, в частности, для германиевых диодов. И наоборот, в кремниевых диодах из-за значительно меньших обратных токов напряжение теплового пробоя получается настолько боль-

шим, что раньше наступает лавинный пробой. Однако это не означает, что в кремниевых диодах не может быть теплового пробоя. Он может происходить при высоких температурах окружающей среды. Кроме того, пробой может начаться как лавинный, а затем, по мере увеличения обратного тока, перейти в тепловой.

В связи с тем что пробивное напряжение при тепловом пробое уменьшается с увеличением теплового сопротивления, на совершенство конструкции диода с точки зрения уменьшения его теплового сопротивления следует обратить особое внимание. Необходимо также отметить, что тепловое сопротивление может увеличиться из-за неправильной установки диода, когда он оказывается теплоизолированным. Напряжение теплового пробоя при этом может существенно уменьшиться. То же самое может произойти при изменении условий в окружающей среде (например, при понижении давления воздуха в связи с подъемом на большую высоту).

Особенности теплового пробоя в реальных диодах

Тепловой пробой в реальных диодах всегда происходит с образованием так называемого «шнура» или канала высокой проводимости, температура в котором превышает среднюю температуру остальной части *p-n*-перехода. В свою очередь, образование шнура может быть вызвано либо дефектами в реальном *p-n*-переходе, либо статистической флуктуацией плотности обратного тока по площади *p-n*-перехода. Действительно, если в каком-нибудь месте *p-n*-перехода в некоторый момент времени плотность обратного тока оказалась несколько выше плотности обратного тока в остальной части *p-n*-перехода, то температура этого места *p-n*-перехода будет выше из-за большей выделяющейся там удельной мощности. Увеличение температуры приведет к росту плотности обратного тока в данном месте *p-n*-перехода в связи с увеличением тепловой генерации носителей либо в самом переходе (см. § 3.10), либо в областях полупроводника, прилегающих к *p-n*-переходу (см. § 3.2, 3.4 и 3.5). Локальное увеличение плотности тока вызовет локальное увеличение температуры, увеличение температуры вызовет увеличение плотности тока и т. д.

Диаметр шнура, возникшего при тепловом пробое, может составлять всего несколько микрометров. Длина его определяется толщиной *p-n*-перехода, т. е. может составлять десятки микрометров. Поэтому, учитывая малый объем шнура, следует заключить, что для развития теплового пробоя в реальных диодах при шнуровании тока требуется очень малая мощность, т. е. тепловой пробой может происходить при малых обратных токах и малых обратных напряжениях. Удельная мощность, выделяе-

мая в едииичном объеме шнура, даже при очень малых обратных токах через диод оказывается довольио большой.

Для возникновения теплового пробоя в соответствии с (3.87) необходим перегрев шнура примерно на 10 К. Это свидетельствует, во-первых, о том, что опять необходима малая мощность для развития теплового пробоя, и, во-вторых, о том, что тепловой пробой — это явление обратимое, если, конечно, ограиичить обратный ток при пробое, не доводя разогрев шнура до появления необратимых физико-химических процессов в полупроводнике.

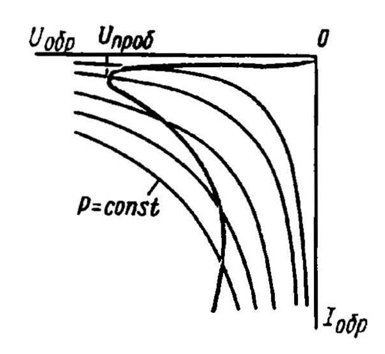


Рис. 3.23. ВАХ при тепловом пробое диода и гиперболы равной мощиости

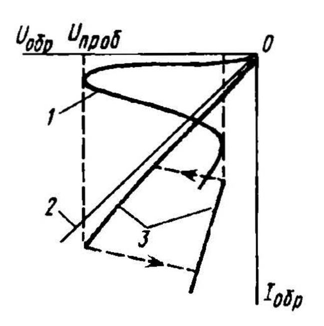


Рис. 3.24. ВАХ шнура (*1*) и шунтирующего сопротивления остальиой части *p-n*-перехода (*2*), которые в сумме могут дать у-образную ВАХ диода (*3*) при тепловом пробое

Следствием малого объема шнура, по которому происходит тепловой пробой, является также малая инерционность процесса теплового пробоя реальных диодов. Тепловые постоянные времени нагрева и охлаждения шнура могут быть порядка 10^{-6}—10^{-7} с. При этом еще надо учесть, что при тепловом пробое напряжение на диоде уменьшается и барьерная емкость *p-n*-перехода диода разряжается через сопротивлеиие шнура с выделением добавочной мощности в шнуре. Это явление способствует ускорению разогрева шнура и уменьшению инерционности процесса теплового пробоя.

Еще одним следствием шнурования тока при тепловом пробое диода является возможность получения своеобразной ВАХ — так называемой у-образной ВАХ, которая на первый взгляд противоречит тепловому механизму пробоя диода. Действительно, с увеличением тока через диод температура *p-n*-перехода все время должна возрастать, что видно как из уравнения теплового баланса (3.78), так и из соотношения (3.79). Таким образом, кривые (гиперболы) равной мощности должны пере-

секать ВАХ диода с учетом тепловыделения в *p-n*-переходе только в одной точке (рис. 3.23). Однако тепловой пробой происходит по шнуру с очень малым поперечным сечением. Если бы удалось из всего *p-n*-перехода выделить шнур, то его ВАХ соответствовала бы всем условиям теплового пробоя (кривая *1* на рис. 3.24). Через остальную часть *p-n*-перехода, площадь которой на несколько порядков больше поперечного сечения шнура, идет обратный ток, т. е. ее можно характеризовать некоторым сопротивлением *R*. Для простоты будем считать сопротивление *R* постоянным — с линейной ВАХ (кривая *2*). Если сопротивление *R* окажется меньше абсолютного значения отрицательного дифференциального сопротивления шнура на падающем участке его ВАХ, то суммарная ВАХ диода (кривая *3*) будет иметь у-образный вид. Следовательно, у-образные ВАХ не противоречат тепловому механизму пробоя диода.

Как отмечалось, пробивное напряжение при тепловом пробое уменьшается с повышением температуры окружающей среды. Однако тепловому пробою может предшествовать лавинный пробой, для которого характерен положительный температурный коэффициент пробивного напряжения. Поэтому температурная зависимость пробивного напряжения для днода при наличии теплового и лавинного пробоев может быть сложной и даже немонотонной, так как при больших температурах тепловой пробой может наступать без предшествующего лавинного пробоя.

§ 3.14. ВЛИЯНИЕ ПОВЕРХНОСТНЫХ СОСТОЯНИЙ НА ВОЛЬТ-АМПЕРНУЮ ХАРАКТЕРИСТИКУ ДИОДА

В реальном полупроводниковом диоде выпрямляющий электрический переход обязательно выходит на поверхность полупроводника. В связи с этим состояние поверхности влияет на ВАХ диода. Это влияние значительно сильнее сказывается на обратной ветви ВАХ, так как обратные токи очень малы.

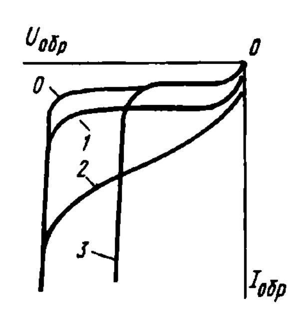


Рис. 3.25. Влияние поверхностных состояний на обратную ветвь ВАХ диода: *0* — без учета поверхностных состояний; *1* — с учетом генерации носителей заряда на поверхности; *2* — при наличии канала поверхностной электропроводности; *3* — при обогащенном слое на поверхности базы

Характер влияния поверхностных состояний зависит от знака и значения поверхностного заряда. Рассмотрим три возможных варианта влияния поверхностных состояний.

Генерация носителей заряда на поверхности полупроводника

Генерация и рекомбинация носителей заряда на поверхности полупроводника, как правило, проходят более интенсивно, чем в его объеме (см. § 1.2). На обратные токи полупроводникового диода генерация носителей заряда на поверхности оказывает такое же влияние, как и генерация носителей в объеме. Однако обратные токи диода зависят от скорости поверхностной генерации, а скорость поверхностной генерации может изменяться со временем из-за изменения поверхностного заряда.

В любом случае обратные токи диода с учетом генерации носителей заряда на поверхности должны быть больше по сравнению с токами без генерации носителей на поверхности (кривая *1* на рис. 3.25).

Каналы поверхностной электропроводности

Если на поверхности полупроводника существует большой поверхностный заряд, совпадающий по знаку с зарядом основных носителей в базе диода, то такой поверхностный заряд отталки-

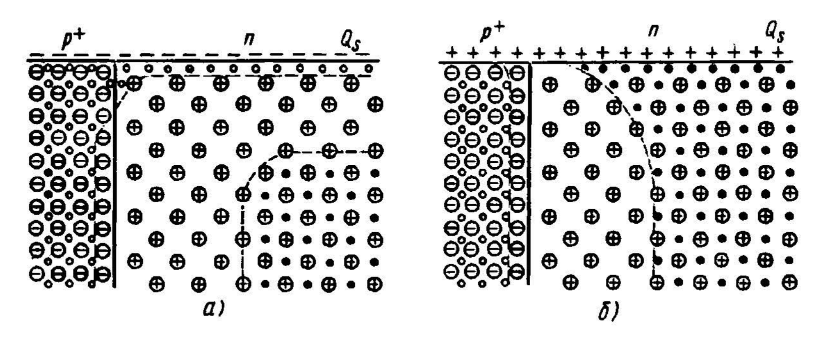


Рис. 3.26. Искажение границ *p-n*-перехода под влиянием поверхностного заряда:

а — при образовании инверсного слоя на поверхности базы диода; *б* — при образовании обогащенного слоя на поверхности базы диода

вает основные носители от поверхности базы и притягивает к поверхности неосновные носители, что приводит к образованию инверсного слоя по всей поверхности базы (рис. 3.26, *а*). При возникновении инверсного слоя увеличивается площадь *p-n*-перехода. Теперь экстракция неосновных носителей из базы будет происходить не только из слоя базы толщиной L_p, примыкающего к *p-n*-переходу в объеме, но и из такого же слоя, примыкающего к поверхности базы (рис. 3.27, *а*). Таким образом,

инверсный слой тоже может принимать участие в сборе неосновных иосителей заряда из базы диода.

В связи с тем что вдоль инверсного слоя проходит ток, напряжение между инверсным слоем и объемом базы ие остается постояииным — по мере удаления от металлургической границы p-n-перехода оно уменьшается (рис. 3.27, *б*). При этом надо учитывать большое удельиое сопротивлеиие инверсиого слоя из-за малой его толщииы. Тогда иа некотором расстоянии l от металлургической граиицы p-n-перехода это напряжение примерно равно kT/q. Следовательио, ие весь инверсный слой участвует в сборе иеосновных носителей, а только его часть протяженностью l. Ту часть инверсиого слоя, которая принимает участие в экстракции неосиовиых носителей заряда, называют *каналом поверхностной электропроводности*.

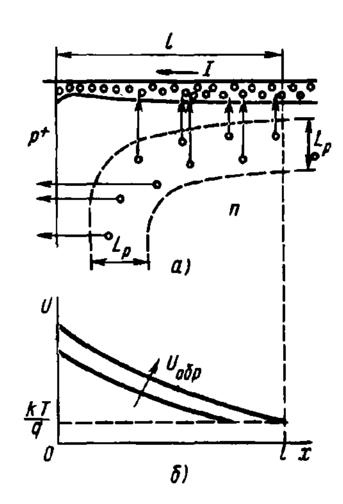


Рис. 3.27. Пути движения неосновных носителей заряда при наличии на поверхности базы канала поверхностной электропроводности (*а*) и распределение напряжения вдоль канала (*б*)

При увеличении обратиого напряжеиия на диоде увеличивается и длина канала поверхиостиой электропроводности (рис. 3.27, *б*). Таким образом, обратная ветвь ВАХ диода при наличии каиала поверхиостиой электропроводности в базовой области ие будет иметь участка насыщения (см. рис. 3.25, кривая *2*).

Поверхиостиый пробой

Поверхностным пробоем p-n-перехода называют пробой перехода, который происходит в месте выхода перехода на поверхность кристалла и иа пробивное напряжение которого оказывают влияние поверхиостиые состояния.

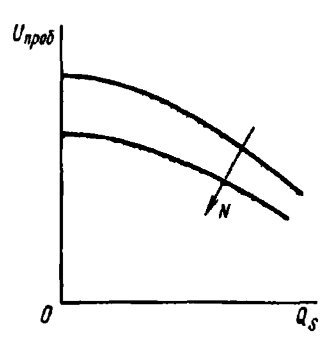


Рис. 3.28. Зависимости пробивного иапряжения диода от поверхностного заряда, создающего обогащенный слой у поверхности базы, при разных концентрациях прнмесей в базе

Если поверхиостиый заряд (заряд поверхиостиых состояиий) имеет знак, противоположиый знаку осиовиых носителей в базе диода, то иа поверхиости базы образуется обогащеииый слой (см. § 1.13). Из-за возникновеиия обогащеиного слоя толщина p-n-перехода у поверхиости базы умеиьшается, так как диффузиоииое электрическое поле перехода проиикает в обогащениый слой иа меиьшую глубину

(см. рис. 3.26, *б*). Влияние поверхностного заряда сказывается в основном на области базы, потому что ее удельное сопротивление велико.

В связи с меньшей толщиной *p-n*-перехода у поверхности пробой диода будет происходить именно там, а пробивное напряжение окажется тем меньше, чем больше сужение перехода у поверхности (см. рис. 3.25, кривая *3*). Таким образом, значение пробивного напряжения зависит в данном случае от плотности поверхностных состояний или от значения поверхностного заряда, создающего обогащенный слой у поверхности базы (рис. 3.28). По своей природе поверхностный пробой может быть лавинным, туннельным или тепловым.

§ 3.15. ПРОЦЕССЫ В ДИОДАХ ПРИ БОЛЬШИХ ПРЯМЫХ ТОКАХ

Прежде чем рассматривать явления в полупроводниковых диодах при больших прямых токах, установим понятия об уровне инжекции. Под уровнем инжекции будем понимать соотношение между концентрацией неосновных носителей и концентрацией основных носителей в состоянии равновесия (или, что почти то же самое, концентрацией примесей). *Низким уровнем инжекции* будем считать такой, при котором концентрация инжектированных неосновных носителей значительно меньше концентрации основных носителей в состоянии равновесия, т. е. для полупроводника *n*-типа

$$p_n - p_{n0} = \Delta p_n \ll n_{n0},$$

для полупроводника *p*-типа

$$n_p - n_{p0} = \Delta n_p \ll p_{p0}.$$

Средним уровнем инжекции будем считать такой, при котором разница между концентрацией инжектированных неосновных носителей и концентрацией основных носителей в состоянии равновесия мала, т. е.

$$\Delta p_n \approx p_n \approx n_{n0}.$$

Высоким уровнем инжекции будем считать такой, при котором

$$\Delta p_n \approx p_n \gg n_{n0}.$$

Прохождение через полупроводниковый диод больших токов, как правило, сопровождается повышением уровня инжекции, по крайней мере в некоторых частях базы диода.

При больших прямых токах в базе полупроводникового диода изменяется концентрация как неосновных, так и основных носителей заряда, появляется электрическое поле, изменяются электрофизические параметры материала базы (время жизни, подвижность или коэффициент диффузии, удельное сопротивле-

ние), изменяются условия как на выпрямляющем, так и на омическом переходе. При больших токах необходимо также учитывать сопротивление базы диода.

Изменение концентрации основных носителей

При инжекции неосновных носителей заряда в базе диода возникает избыточная концентрация этих носителей и нарушается электрическая нейтральность базы. Из-за диффузии в базе и рекомбинации получается неравномерное распределение неосновных носителей в базе (рис. 3.29). Для компенсации инжектированных неосновных носителей заряда через омический переход в базу входят основные носители и также распределяются в базе неравномерно.

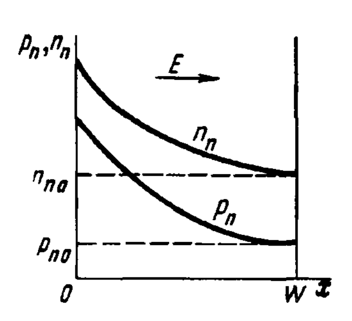


Рис. 3.29. Распределение концентрации неосновных и основных носителей заряда в базе диода при включении его в прямом направлении

С достаточной для практических расчетов точностью можно считать, что при этом восстанавливается электрическая нейтральность базы, т. е. избыточная концентрация основных носителей в базе диода равна избыточной концентрации неосновных носителей:

$$p_n - p_{n0} \approx n_n - n_{n0},$$

или

$$p_n \approx n_n - N_д. \qquad (3.89)$$

Появление электрического поля в базе

Равенство (3.89) является приблизительным, так как для поддержания неравномерного распределения основных носителей в базе необходимо электрическое поле, которое и возникает из-за неполной компенсации заряда неосновных носителей вблизи p-n-перехода.

В принципе нейтрализация заряда инжектированных носителей не может быть абсолютно полной. Действительно, если предположить полную нейтрализацию, то, во-первых, не будет причины, вызывающей неравномерное распределение основных носителей. Во-вторых, неравномерное распределение основных носителей без электрического поля приведет к возникновению диффузионного тока электронов в направлении, противоположном истинному направлению движения электронов в базе диода при его прямом включении. Таким образом, предположение

об абсолютной нейтрализации заряда инжектированных носителей ведет к противоречиям.

Остается принять, что нейтрализация заряда оказывается неполной — избыточная концентрация основных носителей около *p-n*-перехода остается несколько меньшей избыточной концентрации неосновных носителей. Возникающая из-за нескомпенсированного заряда напряженность электрического поля поддерживает неравномерное распределение основных носителей заряда в базе.

Другая составляющая электрического поля в базе диода является следствием прохождения тока через объемное сопротивление базы.

Чтобы определить суммарное значение электрического поля в базе, воспользуемся выражениями для плотности электронного (1.30) и дырочного (1.31) токов, а также выражением для полной плотности тока (1.32) с учетом того, что в установившемся состоянии $\frac{\partial(\varepsilon\varepsilon_0 E)}{\partial t}=0$.

Из условия электрической нейтральности (3.89) для диода с равномерно легированной базой, т. е. при $N_д(x)=\mathrm{const}$, следует, что

$$\mathrm{grad}\, n=\mathrm{grad}\, p.$$

Тогда

$$\mathbf{J}=q(p\mu_p+n\mu_n)\mathbf{E}+q(D_n-D_p)\,\mathrm{grad}\, p$$

и

$$\mathbf{E}=\frac{\mathbf{J}}{q(p\mu_p+n\mu_n)}-\frac{D_n-D_p}{p\mu_p+n\mu_n}\,\mathrm{grad}\, p. \qquad (3.90)$$

Полученный результат показывает, что напряженность электрического поля действительно состоит из двух составляющих. Первая из них равна отношению плотности полного тока к удельной проводимости базы, т. е. представляет собой в соответствии с законом Ома в дифференциальной форме напряженность электрического поля, возникающую из-за прохождения тока через материал базы с конечной проводимостью. Следует отметить, что удельная проводимость в данном случае не является постоянной, она зависит от концентрации носителей заряда и, значит, от плотности тока через диод.

Вторая составляющая напряженности электрического поля определяется градиентом концентрации носителей заряда. Эта составляющая и есть напряженность электрического поля, которое поддерживает неравномерное распределение концентрации основных носителей для обеспечения почти полной электрической нейтральности базы.

Если проинтегрировать (3.90) по координатам, то получим разность потенциалов, которая тоже будет иметь две составляющие. Первая из них — это падение напряжения на объемном

сопротивлении базы. Вторая составляющая представляет собой так называемую ЭДС Дембера, которая была рассмотрена в § 1.12. Эта ЭДС может возникнуть в полупроводнике при неравномерном распределении носителей даже при отсутствии тока через полупроводник, как видно из соотношения (3.90). Так, при освещении полупроводника между освещенными и неосвещенными его частями возникает разность потенциалов. Аналогичную природу имеет послеинжекционная ЭДС (см. § 3.18).

Таким образом, при больших токах в базе диода помимо диффузии носителей заряда необходимо учитывать их дрейф. Кроме того, из-за наличия в базе при больших токах через диод электрического поля напряжение на *p-n*-переходе отличается от напряжения, приложенного к выводам диода.

Изменение электрофизических параметров

Время жизни носителей заряда. При высоком уровне инжекции изменяется ход процесса рекомбинации. Это справедливо для любой модели процесса рекомбинации. Так, при увеличении концентрации неосновных и основных носителей заряда изменяются положения демаркационных уровней и рекомбинация может происходить через такие ловушки, через которые она не происходила при низком уровне инжекции.

Вообще зависимость времени жизни от уровня инжекции определяется многими факторами: типом и расположением энергетических уровней ловушек в запрещенной зоне полупроводника, температурой и т. д. С увеличением уровня инжекции время жизни может уменьшаться, увеличиваться, может наблюдаться и немонотонная зависимость.

Подвижность (коэффициент диффузии). При высоком уровне инжекции может проявляться еще один механизм рассеяния носителей заряда — рассеяние носителей на носителях, что приводит к уменьшению подвижности и коэффициента диффузии носителей заряда. Однако это явление часто можно не принимать во внимание, так как оно начинает сказываться при сравнительно больших концентрациях инжектированных носителей.

Изменение условий на выпрямляющем и омическом переходах. Граничная концентрация неосновных носителей заряда около *p-n*-перехода при больших прямых напряжениях перестает экспоненциально зависеть от напряжения на переходе (см. § 2.3). Можно считать, что при очень больших прямых напряжениях *p-n*-переход полупроводникового диода как бы исчезает и диод становится похожим на резистор с переменным сопротивлением.

В то же время при больших токах нарушается равновесие у омического перехода. Здесь из-за ограниченной скорости движения начинает расти концентрация неосновных носителей заряда, соответственно растет и концентрация основных носителей

(см. § 2.12). База диода при этом заполняется дополнительным количеством электронов и дырок, что приводит к уменьшению ее сопротивления, т. е. *к модуляции сопротивления базы диода.*

§ 3.16. РАСЧЕТ ВОЛЬТ-АМПЕРНОЙ ХАРАКТЕРИСТИКИ ДИОДА ПРИ БОЛЬШИХ ПРЯМЫХ ТОКАХ

Выбор модели диода

Большое количество факторов, которые следует принимать во внимание при расчете ВАХ диода при больших прямых токах, приводит к тому, что в общем виде решить эту задачу не удается.

В связи с этим для выполнения аналитического расчета целесообразно выбрать такую структуру модели полупроводникового диода, которая позволила бы выполнить расчет и в то же время по возможности больше соответствовала бы реальной структуре диода.

Одной из таких структур является полупроводниковый диод с несимметричным электронно-дырочным p^+-n-переходом, с тонкой базой и с идеальным невыпрямляющим контактом к базе. В базе такого диода можно выделить слои с высоким уровнем инжекции (у электронно-дырочного перехода), средним (там, где концентрация неосновных носителей сравнима с концентрацией основных носителей в равновесном состоянии) и с низким уровнем инжекции (вблизи омического перехода) (рис. 3.30).

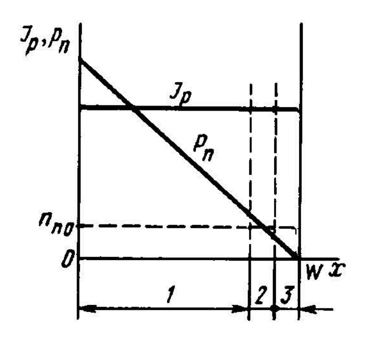


Рис. 3.30. Распределение концентрации неосновных носителей заряда в базе диода с тонкой базой:

1 — слой с высоким уровнем инжекции; *2* — слой со средним уровнем инжекции; *3* — слой с низким уровнем инжекции

Очевидно, что с увеличением тока через диод будет увеличиваться та часть базы, где существует высокий уровень инжекции. При достаточно больших токах почти во всей базе будет высокий уровень инжекции. Поэтому весь расчет можно выполнять только для высокого уровня инжекции. Это первое преимущество выбранной модели для расчета.

Второе преимущество диода с тонкой базой как модели для расчета состоит в том, что ток одного вида носителей (неосновных) во всей базе такого диода остается постоянным. Это следует из того, что во всей тонкой базе практически отсутствует рекомбинация носителей. Она происходит только на омическом переходе.

Третье преимущество рассматриваемой модели следует из несимметрии ее перехода. Если концентрация инжектированных

неосновных носителей у границы p-n-перехода еще не равна полной концентрации основных носителей, что может быть только при очень больших прямых токах, то через несимметричный переход происходит инжекция только одного вида носителей — из сильнолегированной области в слаболегированную. В этом случае весь ток связан с движением носителей одного вида — дырок. Значит, можно рассматривать движение только одного вида носителей во всей базе диода.

Четвертое преимущество рассматриваемой модели заключается в том, что она соответствует реальной структуре большинства полупроводниковых диодов.

Расчет вольт-амперной характеристики диода с тонкой базой при больших прямых токах

Исходя из перечисленных допущений можно считать, что в базе рассматриваемого диода электронный ток равен нулю. Тогда напряженность электрического поля в базе диода может быть найдена из уравнения для плотности электронного тока [см. (1.30)]:

$$J_n = 0 = qn\mu_n E + qD_n \frac{dn_n}{dx}.$$

С учетом соотношения Эйнштейна (1.29)

$$E = -\frac{kT}{q}\frac{1}{n_n}\frac{dn_n}{dx}. \qquad (3.91)$$

Такой способ определения напряженности электрического поля в базе диода несколько отличается от способа, использованного в § 3.15, и в данном случае он более удобен.

Используя условие электрической нейтральности базы (3.89), получим

$$E = -\frac{kT}{q}\frac{1}{p_n + N_{дn}}\frac{dp_n}{dx}. \qquad (3.92)$$

Соотношение (3.92) представляет собой полное значение напряженности электрического поля, включающее поле, связанное с падением напряжения на объемном сопротивлении базы, и поле, обусловленное ЭДС Дембера.

Тогда, подставив (3.92) в (1.31), получим дырочную составляющую плотности тока в базе диода, а следовательно, и ток через диод:

$$J_p = -qD_p \frac{p_n}{p_n + N_{дn}}\frac{dp_n}{dx} - qD_p\frac{dp_n}{dx},$$

или

$$J_p = -qD_p\left(1 + \frac{p_n}{p_n + N_{дn}}\right)\frac{dp_n}{dx}. \qquad (3.93)$$

Выражение (3.93) напоминает формулу для расчета диффузионной составляющей плотности тока (1.27), так как плотность тока получилась пропорциональной градиенту концентрации носителей. Однако такое сходство формально, так как на самом деле в расчете была принята во внимание и напряженность электрического поля в базе диода.

Особенно простой вид принимает выражение для плотности дырочного тока через диод при высоком уровне инжекции, т. е. когда $p_n \gg N_{дn}$:

$$J_p = -q2D_p \frac{dp_n}{dx}. \tag{3.94}$$

Таким образом, дырочный ток через диод прямо пропорционален градиенту концентрации носителей заряда в базе, причем коэффициент пропорциональности не зависит от концентрации носителей. В выражении (3.94) коэффициент диффузии удвоен, что отражает влияние электрического поля в базе диода.

Формальное сходство выражения для плотности тока при высоком уровне инжекции (3.94) с выражением для чисто диффузионного тока (1.27) дает возможность упростить расчет ВАХ диода при высоком уровне инжекции, так как все отличие от рассмотренного ранее расчета при низком уровне инжекции заключается только в удвоении коэффициента диффузии, т. е.

$$J = J_p = \frac{q2D_p p_{n0}}{W_n}\left(\exp\frac{qU_{p\text{-}n}}{kT} - 1\right). \tag{3.95}$$

В уравнении для ВАХ диода при низком уровне инжекции [см. (3.27)] напряжение U имело смысл напряжения на p-n-переходе. Оно же считалось равным напряжению на диоде, так как падением напряжения на базе диода при низком уровне инжекции можно было пренебречь. При высоком уровне инжекции следует учесть разницу в падениях напряжения на переходе и на диоде, пометив индексом p-n напряжение в формуле (3.95). Теперь понятен физический смысл полученной формулы, но необходимо еще найти распределение падений напряжения на диоде.

Расчет падения напряжения на базе в диоде с тонкой базой

Полное падение напряжения на диоде складывается из падения напряжения на p-n-переходе и на базе диода. Падение напряжения на базе диода может быть найдено с помощью соотношения для напряженности электрического поля в базе (3.91) и преобразованием интеграла по координате в интеграл по концентрации носителей заряда:

$$U_Б = \int_0^{U_Б} dU_Б = -\int_{W_n}^{0} E\,dx = \frac{kT}{q}\int_{W_n}^{0} \frac{1}{n_n}\frac{dn_n}{dx}\,dx =$$

$$=\frac{kT}{q}\int\limits_{n_{n0}}^{n_{n\,гр}}\frac{dn_n}{n_n}=\frac{kT}{q}\ln\frac{n_{n\,гр}}{n_{n0}}. \qquad (3.96)$$

Нижний предел у последнего интеграла соответствует концентрации основных носителей в базе в состоянии равновесия. Это можно считать справедливым даже при токах, близких к предельно допустимым для полупроводниковых диодов, при хорошем невыпрямляющем контакте с базой. В этом случае концентрация основных носителей у омического перехода изменяется с изменением тока относительно мало, или, другими словами, мало́ падение напряжения на омическом переходе.

Концентрацию основных носителей около p-n-перехода можно определить из условия электрической нейтральности ($p_{n\,гр}\approx n_{n\,гр}$) и выражения для граничной концентрации неосновных носителей заряда (2.5), т. е. при $x=0$

$$n_{n\,гр}=p_{n\,гр}=p_{n0}\exp\frac{qU_{p\text{-}n}}{kT}. \qquad (3.97)$$

После подстановки (3.97) в (3.96) получаем

$$U_Б=\frac{kT}{q}\ln\left(\frac{p_{n0}}{n_{n0}}\exp\frac{qU_{p\text{-}n}}{kT}\right),$$

или с учетом (1.19)

$$U_Б=\frac{kT}{q}\ln\left(\frac{n_i}{n_{n0}}\right)^2+U_{p\text{-}n}.$$

Тогда полное падение напряжения на диоде

$$U=U_Б+U_{p\text{-}n}=2\frac{kT}{q}\ln\frac{n_i}{n_{n0}}+2U_{p\text{-}n},$$

отсюда

$$U_{p\text{-}n}=\frac{U}{2}-\frac{kT}{q}\ln\frac{n_i}{n_{n0}}=\frac{U}{2}+\frac{kT}{q}\ln\frac{n_{n0}}{n_i}. \qquad (3.98)$$

Подставив последнее выражение (3.98) в уравнение ВАХ (3.95), получим

$$J=\frac{q2D_p p_{n0}}{W_n}\left[\left(\exp\frac{qU}{2kT}\right)\frac{n_{n0}}{n_i}-1\right].$$

Единицей в квадратных скобках можно пренебречь, так как $n_{n0}\gg n_i$ и $\exp\frac{qU}{2kT}\gg 1$. Тогда

$$J=\frac{q2Dn_i}{W}\exp\frac{qU}{2kT}. \qquad (3.99)$$

Здесь, как и ранее, опущены индексы у параметров материала базы диода (все входящие в формулу величины относятся к слаболегированной области, т. е. к базе диода).

Полученное выражение напоминает уравнение ВАХ диода с тонкой базой при низком уровне инжекции (3.27) с учетом (3.31), но есть и некоторые отличия:

1) вместо обычного коэффициента диффузии стоит его удвоенное значение, что отражает влияние электрического поля в базе;

2) в показателе экспоненты стоит уменьшенное вдвое напряжение, что связано с распределением приложенного к диоду напряжения приблизительно пополам между *p-n*-переходом и базой диода;

3) вместо концентрации неосновных носителей в последней формуле стоит собственная концентрация. Значит, вид ВАХ при высоком уровне инжекции перестает зависеть от концентрации примесей в слаболегированной области (в базе диода). Объясняется это тем, что при высоком уровне инжекции свойства полупроводника базы определяются в основном концентрацией инжектированных носителей. Полупроводник базы ведет себя при этом аналогично собственному полупроводнику.

Диапазон применимости полученного соотношения для вольт-амперной характеристики диода при больших токах

Диапазон применимости соотношения (3.99) ограничен снизу тем, что при малых токах через диод не выполняется условие высокого уровня инжекции; сверху — тем, что при очень больших токах становятся несправедливыми принятые условия на границах *p-n*- и омического переходов с базой.

Нижняя граница диапазона применимости. Строго нижнюю границу диапазона применимости соотношения (3.99) следовало бы определить по концентрации носителей заряда в соответствии с определением уровня инжекции. ВАХ, рассчитанная по приближенным формулам для низкого и высокого уровней инжекции, должна быть непрерывной. Тогда в качестве предельного тока, являющегося границей диапазонов применимости приближений, можно взять такой, при котором вычисленные токи в том и другом приближении окажутся одинаковыми, т. е. соотношения (3.27) с учетом (3.31) и (3.99) имеют вид

$$\frac{qD_p p_{n0}}{W_n}\left(\exp\frac{qU}{kT}-1\right)=\frac{q2D_p n_i}{W_n}\exp\frac{qU}{2kT}. \qquad (3.100)$$

Если пренебречь единицей в формуле тока при низком уровне инжекции, то

$$\exp\frac{qU}{2kT}=\frac{2n_i}{p_{n0}}.$$

Подставив этот результат в любую часть выражения (3.100), получим для граничной плотиости тока

$$J_н = \frac{q2D_p n_i 2n_i}{W_n p_{n0}},$$

или после преобразований с помощью соотношения (1.19)

$$J_н = \frac{4qDN}{W}. \quad (3.101)$$

Здесь опять опущены индексы, которые должны соответствовать слаболегированной области диода, т. е. его базе.

Оценка граничной плотности тока с помощью соотношения (3.101) показывает, что $J_н \approx 0{,}1$ А/см^2. Это примерно на три порядка меньше, чем предельно допустимая плотность тока диода. Следовательно, практически на всей прямой ветви ВАХ полупроводниковый диод работает в условиях больших токов.

Верхняя граница диапазона применимости. Верхняя граница диапазона применимости соотношения (3.99) определяется неравенством $U_{p\text{-}n} < \varphi_{кон}$.

Подставив в это неравенство выражения (3.98) и (2.2), получим

$$\frac{U}{2} - \frac{kT}{q}\ln\frac{n_i}{n_{n0}} < \frac{kT}{q}\ln\frac{p_{p0}n_{n0}}{n_i^2},$$

или

$$\frac{U}{2} < \frac{kT}{q}\ln\frac{p_{p0}}{n_i}.$$

Следовательно,

$$J < \frac{2qD_p n_i}{W_n}\frac{p_{n0}}{n_i},$$

или

$$J_в = \frac{2qDN^+}{W}, \quad (3.102)$$

где N^+ — концентрация примесей в сильнолегированной области, а остальные величины следует брать для слаболегированной области диода.

Порядок плотности тока, соответствующий верхней границе диапазона применимости, вычисленной по соотношению (3.102), около 100 А/см^2, что совпадает с предельно допустимой плотностью тока полупроводниковых диодов.

Вольт-амперная характеристика диода с толстой базой при больших токах

В диодах с толстой базой, когда толщина базы во много раз превышает диффузионную длину неосновных носителей заряда (например, в диодах на основе материала с большой шириной запрещенной зоны), сопротивление базы можно считать не зависящим от тока, т. е. можно пренебречь модуляцией сопротивления базы. Тогда при условии экспоненциальной зависимости граничной концентрации неосновных носителей около перехода от напряжения на нем уравнение ВАХ диода с толстой базой может быть представлено в виде

$$I = I_{нас}\left[\exp\frac{q(U - IR_Б)}{kT} - 1\right],$$

где $R_Б$ — сопротивление базы диода.

Если сопротивление базы известно, то эта формула позволяет рассчитать ВАХ, принимая в качестве независимой переменной значение тока.

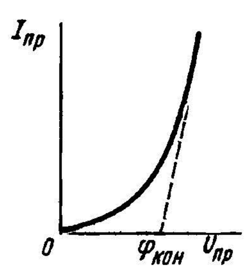


Рис. 3.31. Прямая ветвь диода с толстой базой

При больших токах, когда напряжение на *p-n*-переходе приближается к контактной разности потенциалов, это напряжение как бы стабилизируется. Тогда ВАХ диода с толстой базой принимает вид

$$U = \varphi_{кон} + IR_Б.$$

Таким образом, при больших токах на ВАХ диода с толстой базой может наблюдаться прямолинейный участок (рис. 3.31). Это дает возможность оценить контактную разность потенциалов на *p-n*-переходе диода.

Однако в большинстве случаев сопротивление базы не является постоянным при изменении тока и расчет ВАХ сильно усложняется. Но заранее можно сказать, что при напряжениях на переходе, близких к контактной разности потенциалов, ВАХ вернее всего, не будет экспоненциальной, так как экспоненциальный ее вид прежде всего связан с экспоненциальной зависимостью концентрации инжектированных носителей заряда от напряжения.

§ 3.17. ВОЛЬТ-АМПЕРНАЯ ХАРАКТЕРИСТИКА ДИОДА В ПОЛУЛОГАРИФМИЧЕСКИХ КООРДИНАТАХ

На прямой ветви ВАХ диода, построенной в полулогарифмических координатах, могут существовать участки, связанные с различными механизмами образования тока. Там, где ВАХ экспонен-

циальна, в даиной системе коордииат получается прямолииейный отрезок (рис. 3.32).

При напряжеиии на диоде, равном нулю, ток равен нулю. Следовательно, $\lg I \to -\infty$.

При малых напряжениях на диоде ток может определяться рекомбинацней в *p-n*-переходе. Тогда он зависит от напряжеиия как $\exp\frac{qU}{2kT}$ [см. (3.58)], что дает на графике отрезок прямой линии с угловым коэффициентом 1/2.

При бóльших напряжениях преобладает ток, связаниый с инжекцией носителей заряда и зависящий от напряжеиия как $\exp\frac{qU}{kT}$ [см. (3.27)], что дает на графике отрезок прямой линии с угловым коэффициентом 1.

При дальнейшем увеличеиии тока наряду с инжекцией носителей заряда начинает сказываться модуляция сопротивления базы, что опять дает зависимость $\exp\frac{qU}{2kT}$ [см. (3.99)] или на графике — отрезок прямой с угловым коэффициентом 1/2.

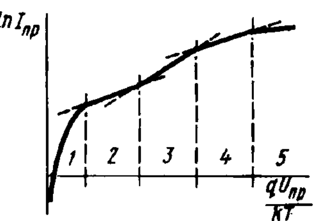


Рис. 3.32. Прямая ветвь ВАХ диода:

1 — участок малых токов; *2* — участок с преобладаинем рекомбииации иосителей в *p-n*-переходе; *3* — участок с преобладанием процесса инжекции; *4* — участок, связаниый с модуляцией сопротивлеиия базы при высоком уровие инжекции; *5* — участок, связанный с нарушеинем условий иа границах базы с *p-n*-переходом и с омическим переходом

Наконец, при очень больших прямых токах ВАХ диода становится неэкспоненциальной из-за нарушения условий на *p-n*-переходе и на омическом переходе. Соответствующий отрезок на графике не будет прямой линией.

Таким образом, построение ВАХ диода в полулогарифмическом масштабе может дать информацию относительно возможных механизмов прохождения тока. Надо учесть, что прямолинейные участки на этой кривой могут плавно переходнть один в другой, так что определить их границы точно не представляется возможным. Иногда некоторые участки отсутствуют. Так, для некоторых кремниевых диодов участок, соответствующий рекомбинационному току, может сразу перейти в участок, соответствующий высокому уровню инжекции. Их граиицу обнаружить нельзя. Искажает вид характеристик также измененне времени жизни иосителей, температуры, наличие каналов поверхностной электропроводности и т. п.

§ 3.18. ПЕРЕХОДНЫЕ ПРОЦЕССЫ В ДИОДАХ

Переходные процессы в полупроводниковых диодах связаиы в основном с двумя явлениями, происходящими при быстром изменении напряжения на диоде или тока через диод.

Первое из них — это накопление неосновных носителей заряда в базе диода при его прямом включении и их рассасывание при уменьшении напряжения. Так как электрическое поле в базе диода обычно невелико, то движение неосновных носителей в базе определяется законами диффузии и происходит относительно медленно. Поэтому накопление носителей в базе и их рассасывание могут влиять на свойства диодов в режиме переключения.

Второе явление, происходящее в диодах при их переключении, — это перезарядка барьерной емкости, что также происходит не мгновенно и поэтому может влиять на свойства диодов.

При сравнительно больших плотностях прямого тока через диод существенно накопление неосновных носителей в базе диода, а перезарядка барьерной емкости диода является второстепенным процессом. При малых плотностях тока переходные процессы в диоде определяются перезарядкой барьерной емкости диода, накопление же неосновных носителей заряда в базе практически не сказывается.

Временны́е зависимости напряжения и тока, характеризующие переходные процессы в полупроводниковом диоде, зависят также от сопротивления внешней цепи, в которую включен диод. Поэтому рассмотрим четыре предельных случая переходных процессов в полупроводниковом диоде с несимметричным p^+-n-переходом.

Процессы при больших напряжениях и токах

Работа диода в схеме с генератором напряжения. Рассмотрим процессы, происходящие в полупроводниковом диоде при его включении на генератор напряжения, т. е. при включении диода в цепь с малым полным сопротивлением (в том числе и с малым сопротивлением источника питания) по сравнению с сопротивлением диода.

При подаче на диод прямого напряжения ток через диод устанавливается не сразу, так как с течением времени происходит накопление в базе инжектированных через p-n-переход неосновных носителей (дырок) и уменьшение в связи с этим сопротивления базы (рис. 3.33). Этот процесс *модуляции сопротивления базы* происходит не мгновенно, так как накопление дырок в базе диода связано с относительно медленным процессом диффузии их от p-n-перехода в глубь базы.

По мере накопления дырок и уменьшения сопротивления базы происходит перераспределение всего внешнего напряжения между сопротивлением базы и p-n-переходом; падение напряжения на базе диода уменьшается (рис. 3.33, *б*), а на p-n-переходе увеличивается (рис. 3.33, *в*), вызывая увеличение уровня инжекции (рис. 3.33, *д*). При длительном прохождении прямого тока процесс инжекции дырок уравновешивается процессом их рекомбинации.

Возникает некоторое установившееся состояние, характеризующееся таким распределением дырок в базе, при котором их концентрация превышает равновесную вблизи *p-n*-перехода и снижается, стремясь к равновесной, при удалении от него в глубь базы (кривая *4* на рис. 3.33, *д*).

О значении тока через *p-n*-переход можно судить по градиенту концентрации дырок в базе диода около *p-n*-перехода (см. § 3.4). Градиент концентрации дырок около *p-n*-перехода возрастает

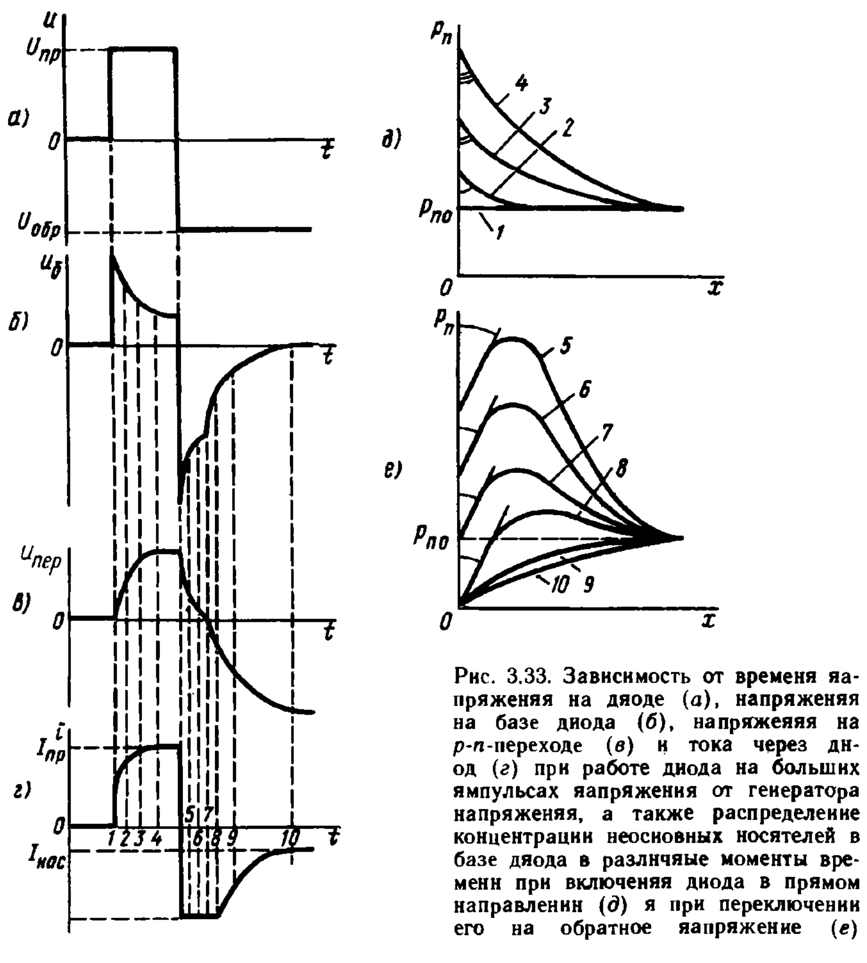


Рис. 3.33. Зависимость от времени напряжения на диоде (*а*), напряжения на базе диода (*б*), напряжения на *p-n*-переходе (*в*) и тока через диод (*г*) при работе диода на больших импульсах напряжения от генератора напряжения, а также распределение концентрации неосновных носителей в базе диода в различные моменты времени при включения диода в прямом направлении (*д*) и при переключении его на обратное напряжение (*е*)

с увеличением напряжения на *p-n*-переходе при постоянном напряжении на диоде и при уменьшении напряжения на базе диода (рис. 3.33, *д*). Следует отметить, что сопротивление базы диода уменьшается не только из-за увеличения концентрации неосновных носителей (дырок), но и из-за увеличения концентрации основных носителей (электронов). Концентрация электронов около *p-n*-перехода возрастает в соответствии с принципом электрической ней-

тральиости, согласио которому в любой части базовой области сумма всех зарядов должна быть равна нулю. Очевидио, что число накопленных дырок в базе тем больше, чем больше ток через диод и чем больше время жизии дырок. Кроме того, число накоплеииых дырок зависит от геометрии базы.

При переключении диода с прямого напряжеиия на обратиое в иачальный момент наблюдается большой обратный ток, ограиичеииый в осиовном иоследовательиым сопротивлеиием базы диода. Источиик питаиия вместе с сопротивлением базы в это время можно считать генератором тока для *p-n*-перехода.

После переключения диода на обратиое напряжение начинается процесс рассасывания неосновных иосителей, накопленных в базе. Из-за ограничения обратиого тока коицеитрация дырок в базе около *p-n*-перехода не может мгновенно уменьшиться до равновесного значения. Пока коицеитрация дырок в базе около *p-n*-перехода превышает равновесное значение, иа *p-n*-переходе сохраняется прямое падеиие напряжения (рис. 3.33, *в*).

После умеиьшения концентрации дырок в базе около *p-n*-перехода до нуля обратный ток начииает уменьшаться со временем, о чем свидетельствует уменьшение градиеита концентрации дырок около *p-n*-перехода (рис. 3.33, *е*). С течеиием времеии все накопленные в базе дырки уходят через *p-n*-переход или рекомбииируют в базе днода, в результате чего обратный ток умеиьшается до стационарного значеиия тока иасыщения (рис. 3.33, *г*). В это время заканчивается восстановлеиие обратиого сопротивления диода.

Процесс рассасывания накопленных носителей происходит значи-

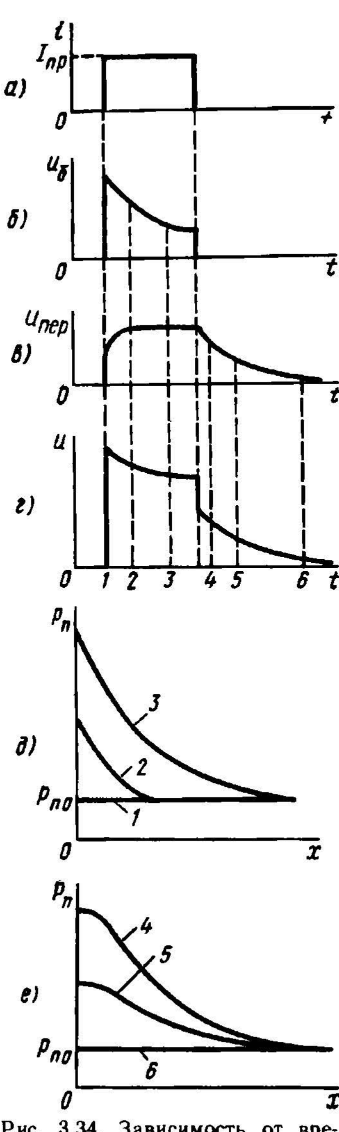


Рис. 3.34. Зависимость от времени тока через диод (*а*), напряжения на базе (*б*), напряжения на *p-n*-переходе (*в*) и напряжения на диоде (*г*) при работе диода на больших импульсах тока в схеме с генератором тока, а также распределение концентрации иеосиовных носителей заряда в базе диода в различные моменты времени при включении диода (*д*) и при выключении диода (*е*)

тельно медлениее процесса их накоплеиия, поэтому имеиио процесс рассасываиия и определяет частотиые свойства большинства диодов.

Работа диода в схеме с генератором тока соответствует включению диода в схему с большим сопротивлением, которое и определяет значение тока в цепи с диодом. Рассмотрим процессы, происходящие в диоде, при прохождеиии через диод импульса прямого тока прямоугольиой формы (рис. 3.34).

В первый момент прохождения через диод импульса прямого тока на диоде падает отиосительно большое напряжеиие, которое в дальиейшем умеиьшается из-за уменьшения сопротивлеиия базовой области диода, связанного с накоплением иеравиовесиых носителей в базе.

После окончания процесса накопления неосновиых иосителей в базе напряжения на диоде, иа базе диода и на *p-n*-переходе достигают установившихся значений. Распределеиие дырок в базе в это время также соответствует некоторому устаиовившемуся состояиию (кривая *3* на рис. 3.34, *д*).

В момент окончания импульса тока через диод, т. е. в момент разрыва цепи с диодом исчезает падеиие иапряжения на объемном сопротивлеиии базы диода (рис. 3.34, *б*). Концентрация дырок в базе около *p-n*-перехода мгновенно измеииться не может. Поэтому напряжение иа *p-n*-переходе и соответственно иа диоде после выключения тока уменьшается замедленно по мере рекомбннации неравновесных носителей в базе. Изменения в распределении концентрации дырок в базе диода со времеием показаиы иа рис. 3.34, *е*. Остаточное иапряжение на диоде уменьшится до нуля после рекомбинации всех неравиовесных носителей заряда в базовой области диода.

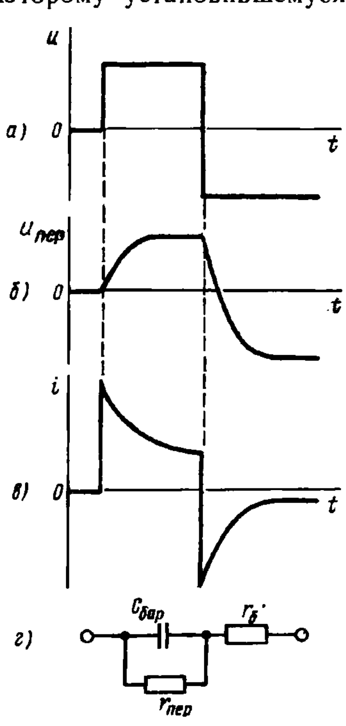


Рис. 3.35. Зависимость от времени напряжения иа диоде (*а*), напряжения на *p-n*-переходе (*б*) и тока через диод (*в*) при малых импульсах напряжения в схеме с генератором иапряжения, а также эквивалентная схема диода для малых сигналов (*г*)

Процессы при малых напряжениях и токах

Работа диода в схеме с генератором напряжения. При приложении к диоду малого прямого напряжения (рис. 3.35) эффект модуляции сопротивления базы диода из-за малого уровня инжекции пренебрежимо мал. Поэтому сопротивление диода в данном случае имеет емкостный

характер. В первый момент напряжение на p-n-переходе близко к нулю, а ток через диод ограничен только сопротивлением базы диода (рис. 3.35). По мере заряда барьерной емкости напряжение на p-n-переходе и ток через диод стремятся к некоторым установившимся значениям, которые определяются активной составляющей сопротивления p-n-перехода.

В момент переключения диода напряжение на барьерной емкости не может измениться мгновенно, оно достигает установившегося значения через некоторое время. Ток через диод также зависит от времени, что характерно для емкостного сопротивления.

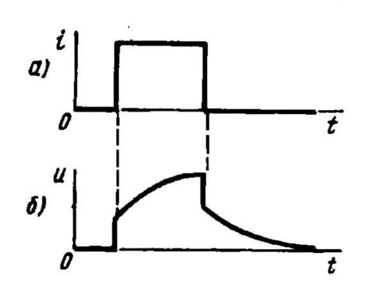


Рис. 3.36. Зависимость тока через диод (*а*) и напряжения на диоде (*б*) при работе диода на малых импульсах тока в схеме с генератором тока

Работа диода в схеме с генератором тока. Осциллограммы импульса прямого тока малой амплитуды и падения напряжения на диоде при питании диода от генератора тока представлены на рис. 3.36. В первый момент пропускания импульса тока через диод весь ток состоит из емкостной составляющей. Поэтому напряжение на диоде в первый момент определяется падением напряжения на сопротивлении базы диода. По мере заряда барьерной емкости увеличивается и напряжение на диоде.

При выключении диода на нем некоторое время сохраняется остаточное напряжение, уменьшающееся со временем. Остаточное напряжение в данном случае связано с тем, что барьерная емкость остается еще заряженной. По мере разряда этой емкости через активное сопротивление p-n-перехода диода уменьшается напряжение на емкости и остаточное напряжение на диоде.

§ 3.19. ВЫПРЯМИТЕЛЬНЫЕ ПЛОСКОСТНЫЕ НИЗКОЧАСТОТНЫЕ ДИОДЫ

Выпрямительный диод — это полупроводниковый диод, предназначенный для преобразования переменного тока в постоянный.

Выпрямительные плоскостные низкочастотные диоды обычно используют для выпрямления переменного тока промышленной частоты (50 Гц). В бортовой аппаратуре частота переменного напряжения составляет 400 Гц. Значительно реже выпрямительные низкочастотные диоды должны работать при еще больших частотах. Так, в транзисторных преобразователях напряжения частота переменного тока, выпрямляемого диодом, достигает нескольких десятков килогерц.

Из основных и справочных параметров выпрямительных диодов необходимо отметить:

1) максимально допустимый прямой ток $I_{пр\,max}$;

2) прямое напряжение на диоде при заданном значении прямого тока $U_{пр}$; обычно прямое напряжение иа диоде указывается при максимально допустимом прямом токе через диод;

3) максимально допустимое обратное напряжение $U_{обр\,max}$; оно обычно значительно меньше пробивного;

4) обратный ток при заданиом обратном иапряжении $I_{обр}$; обычно обратный ток указывается при максимально допустимом обратном напряжении;

5) диапазон рабочих температур окружающей среды.

В зависимости от значения максимально допустимого прямого тока выпрямительные диоды подразделяют на диоды малой мощности (прямой ток до 0,3 А), средней мощности (прямой ток от 0,3 до 10 А) и большой мощности (прямой ток более 10 А).

При производстве первых выпрямительных плоскостных диодов (50-е годы нашего века) в качестве исходного полупроводникового материала использовали германий, технология получения и очистки монокристаллов которого к этому времеии была уже освоена. Значительно позже было налажено производство кремниевых выпрямительных плоскостных диодов. В связи с существенными преимуществами кремниевых выпрямительных плоскостных диодов они практически целиком вытеснили германиевые выпрямительные плоскостные диоды из массового производства. В последние годы налажен выпуск выпрямительных плоскостных диодов из арсенида галлия.

Кремниевые диоды

Технология изготовления и конструкция. Большинство из разнообразных типов кремниевых выпрямительных плоскостных диодов с технологической точки зрения являются диффузионными. Электронно-дырочные переходы таких диодов формируют диффузией алюминия или бора в кристаллы кремния с электропроводностью n-типа и диффузией фосфора в кристаллы кремния с электропроводностью p-типа. В выпрямительных диодах старых разработок для формирования p-n-переходов используют вплавление алюминия в кристаллы кремния с электропроводностью n-типа, а в кристаллы кремния p-типа — вплавление сплава олова с фосфором или золота с сурьмой.

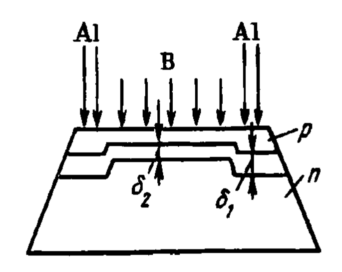


Рис. 3.37. Структура лавинного диода, p-n-переход которого сформирован диффузией бора в центральной части кристалла и диффузией алюминия в кольцевой части

Необходимую площадь p-n-перехода рассчитывают, исходя из значения допустимого прямого тока диода и учитывая допустимую плотность прямого тока, которая для кремниевых p-n-переходов равна 200 А/см^2. Толщина исходных кристаллов кремния составляет 0,2...0,4 мм.

Пробой реальных p-n-**переходов** часто происходит вблизи поверхности полупроводника, т. е. в местах выхода перехода на поверхность кристалла (см. § 3.14). Для исключения возможности поверхностного пробоя толщину диффузионного перехода (δ_1) в месте выхода его на поверхность кристалла кремния, т. е. в кольцевой части по периферии кристалла, желательно сделать большей по сравнению с толщиной перехода (δ_2) в центральной части кристалла (рис. 3.37). Достичь этого можно двумя способами. Во-первых, кольцевую часть перехода можно сформировать диффузией алюминия, а центральную часть — диффузией бора. Коэффициент диффузии алюминия больше коэффициента диффузии бора, а растворимость алюминия в кремнии меньше, чем бора. Поэтому градиент концентрации акцепторной примеси в кольцевой части оказывается меньше, чем в центральной части перехода. Толщина диффузионного перехода зависит от градиента концентрации примеси и будет больше в кольцевой части перехода (см. рис. 3.17, *б*). Пробивное напряжение кольцевой части перехода оказывается при этом больше пробивного напряжения центральной его части. Пробивное напряжение всего диода не зависит от состояния поверхности кристалла; из-за меньшего количества дефектов в центральной части перехода плотность обратного тока при лавинном пробое будет равномерно распределена по всей площади. Такие диоды, называемые *лавинными*, могут выдерживать значительные импульсные перегрузки по обратному току.

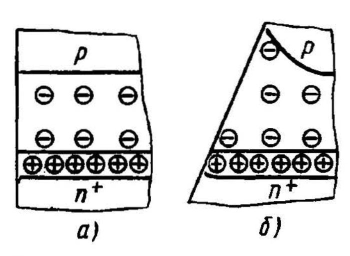


Рис. 3.38. Структура p-n-перехода в месте выхода его на поверхность кристалла без фаски (*а*) и с положительной фаской (*б*), которая обеспечивает увеличение толщины p-n-перехода у поверхности кристалла

Во-вторых, уменьшить вероятность поверхностного пробоя можно путем снятия фаски по периметру кристалла кремния после создания в нем p-n-перехода. Для уменьшения напряженности электрического поля в месте выхода перехода на поверхность кристалла и соответственно для увеличения пробивного напряжения поверхностного пробоя надо путем шлифовки краев кристалла снять фаску от сильнолегированной области к слаболегированной (прямая или положительная фаска). При такой конфигурации фаски толщина p-n-перехода вблизи поверхности кристалла увеличится, что вызвано сохранением электрической нейтральности p-n-перехода (рис. 3.38).

Частотные свойства выпрямительных плоскостных диодов с выпрямляющим переходом в виде p-n-перехода, работающих обычно при большом уровне инжекции, определяются процессами накопления и рассасывания неосновных носителей заряда в базе (см. § 3.18). Поэтому для улучшения частотных свойств кремниевых плоскостных диодов в исходные кристаллы кремния про-

водят диффузию золота, примесь которого создает энергетические уровни рекомбинационных ловушек и уменьшает время жизни неосновных носителей заряда. Такие кремниевые плоскостные диоды иногда называют *частотными*, подчеркивая тем самым, что они способны работать при повышенных частотах (до 100 кГц).

Для защиты от внешних воздействий и обеспечения хорошего теплоотвода кристалл с *p-n*-переходом монтируют в корпусе. Диоды малой мощности обычно оформляют в пластмассовом корпусе с гибкими внешними выводами, диоды средней мощности — в металлостеклянном корпусе с жесткими внешними выводами, диоды большой мощности — в металлостеклянном или металлокерамическом корпусе, т. е. со стеклянным или керамическим проходным изолятором. Пример конструктивного оформления выпрямительного диода показан на рис. 3.39. Кристалл кремния с *p-n*-переходом припаивают к медному основанию (к ножке корпуса) таким образом, чтобы получить омический переход между металлом и полупроводником. Для этого используют припой с примесью донора, если кристалл припаивают к основанию *n*-областью, и с примесью акцептора, если *p*-областью. Наличие в припое химического элемента, создающего в запрещенной зоне кремния уровни рекомбинационных ловушек (золото), обеспечивает уменьшение времени жизни неосновных носителей заряда вблизи омического перехода и ускорение процесса рассасывания накопленных неосновных носителей. В мощных выпрямительных диодах с большой площадью кристалла между медным основанием и кристаллом кремния помещают термокомпенсирующую прокладку из вольфрама или ковара, имеющую примерно тот же коэффициент линейного расширения, что и кремний. Таким образом устраняют или существенно уменьшают механические напряжения в кристалле кремния, возникающие при изменении температуры.

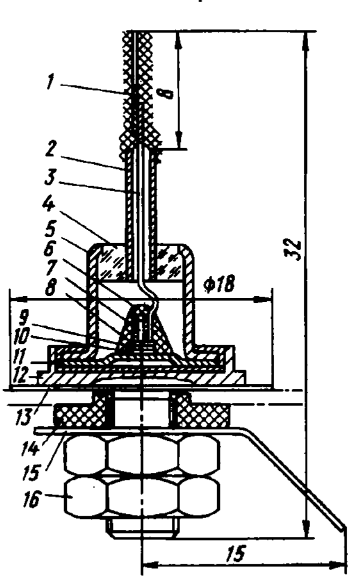


Рис. 3.39. Конструкция одного из кремниевых плоскостных диодов:
1 — наружный вывод; *2* — трубка; *3* — внутренний вывод; *4* — проходной изолятор; *5* — корпус; *6* — защитное покрытие; *7* — верхний электрод на кристалле полупроводника; *8* — кристалл кремния с *p-n*-переходом; *9* — базовая область диодной структуры; *10* — слой припоя; *11* — кристаллодержатель; *12* — ножка корпуса; *13* — слюдяная изолирующая шайба; *14* — изоляционная втулка; *15* — токоотводящий лепесток; *16* — крепежные гайки

Электрические свойства. Наибольшую информацию об электрических свойствах выпрямительных диодов можно получить из ВАХ. Вольт-амперные характеристики одного из выпрямительных

кремниевых плоскостных диодов при разных температурах окружающей среды приведены на рис. 3.40.

Максимально допустимые прямые токи кремниевых плоскостных диодов различных типов составляют 0,1...1600 А. Падение напряжения на диодах при этих токах не превышает обычно 1,5 В. С увеличением температуры прямое напряжение уменьшается, что связано с уменьшением высоты потенциального барьера *p-n*-перехода и с перераспределением носителей заряда по энергиям (см. § 3.2).

Обратная ветвь ВАХ кремниевых диодов не имеет участка насыщения обратного тока, так как обратный ток в кремниевых диодах вызван процессом генерации носителей заряда в *p-n*-переходе (см. § 3.10). Пробой кремниевых диодов имеет лавинный

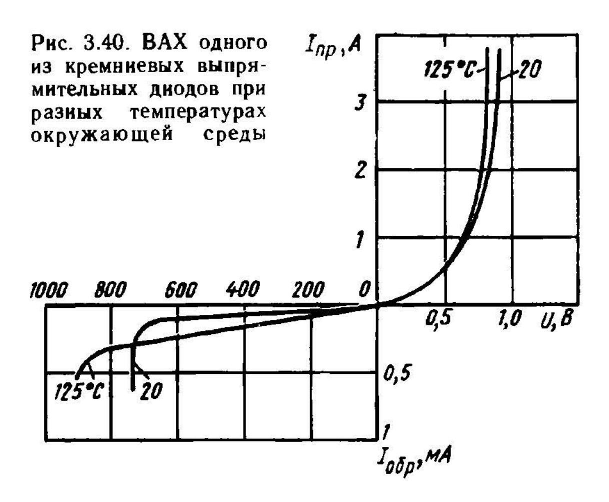


Рис. 3.40. ВАХ одного из кремниевых выпрямительных диодов при разных температурах окружающей среды

характер. Поэтому пробивное напряжение с увеличением температуры увеличивается (см. § 3.11). Для некоторых типов кремниевых диодов при комнатной температуре пробивное напряжение может составлять 1500...2000 В. Максимально допустимое обратное напряжение — это напряжение, при котором гарантируется надежная работа диода. Поэтому максимально допустимое обратное напряжение, значение которого указывается в справочниках, значительно меньше пробивного напряжения. Обычно выбирают $U_{обр\,max} = (0{,}5...0{,}8)U_{проб}$.

Для преобразования переменного тока в постоянный в схемах с напряжением, превышающим максимально допустимое обратное напряжение отдельного диода, промышленностью выпускаются выпрямительные столбы. *Выпрямительный полупроводниковый столб* — это совокупность выпрямительных полупроводниковых диодов, соединенных последовательно и собранных в единую конструкцию, имеющую два вывода. Максимально допустимое обратное напряжение кремниевых выпрямительных столбов составляет несколько киловольт.

Для удобства применения выпрямительных диодов в выпрямителях, собранных по различным мостовым схемам (однофазным, трехфазным), а также в выпрямителях с удвоением напряжения промышленностью выпускаются выпрямительные полупроводниковые блоки. *Выпрямительный полупроводниковый блок* — это полупроводниковый блок, собранный из выпрямительных полупроводниковых диодов, соединенных по определенной электрической схеме и оформленных в единую конструкцию, имеющую более двух выводов.

Диапазон рабочих температур для кремниевых выпрямительных диодов ограничен значениями —60... + 125°С. Нижний предел рабочих температур обусловлен различием температурных коэффициентов линейного расширения различных элементов конструкции диода: при низких температурах возникают механические напряжения, которые могут привести к растрескиванию полупроводникового кристалла. При необходимости этот предел рабочих температур может быть существенно уменьшен, т. е. сдвинут в область более низких температур. Принципиальное ограничение при этом может быть связано с энергией ионизации примесей в различных областях диодной структуры. Но энергия ионизации примесей в кремнии, которые обеспечивают электропроводность *p*- и *n*-типов, мала. Поэтому уже при температуре в несколько десятков кельвин все акцепторы и доноры оказываются ионизированными. С уменьшением температуры необходимо учитывать также увеличение прямого напряжения на диоде, которое происходит из-за увеличения высоты потенциального барьера на *p-n*-переходе.

Верхний предел диапазона рабочих температур выпрямительных диодов определяется резким ухудшением выпрямления в связи с ростом обратного тока — сказывается тепловая генерация носителей заряда в результате ионизации атомов полупроводника. Таким образом, верхний предел диапазона рабочих температур выпрямительных кремниевых диодов, как и большинства других полупроводниковых приборов, связан с шириной запрещенной зоны исходного полупроводникового материала.

Германиевые диоды

Технология изготовления и конструкция. Первой технологией формирования плоскостных *p-n*-переходов в массовом производстве полупроводникоаых приборов была технология вплавления примесей. Поэтому все германиевые плоскостные диоды — это сплавные диоды старых разработок.

Для создания сплавного *p-n*-перехода в германиевом монокристалле используют кристаллы германия с электропроводностью *n*-типа толщиной 0,2...0,4 мм. В качестве акцепторной примеси используют индий, который вплавляют в кристалл германия при температуре 500...550°С. Площадь перехода определяется необходимым значением максимально допустимого прямо-

го тока диода и максимально допустимой плотностью прямого тока, которая для германиевых *p-n*-переходов равна 100 А/см^2.

Электрические свойства. ВАХ одного из германиевых плоскостных диодов при разных температурах окружающей среды приведены на рис. 3.41. Основные отличия ВАХ и параметров германиевых и кремниевых выпрямительных плоскостных диодов заключаются в следующем.

1. Прямое напряжение на германиевом диоде при максимально допустимом прямом токе приблизительно в два раза меньше, чем на кремниевом диоде. Это вызвано меньшей высотой потенциального барьера германиевого *p-n*-перехода является существенным, но, к сожалению, единственным преимуществом перед кремниевыми выпрямительными диодами.

2. Существование тока насыщения при обратном включении германиевого диода, что вызвано механизмом образования обратного тока — процессом экстракции неосновных носителей заряда из прилегающих к *p-n*-переходу областей.

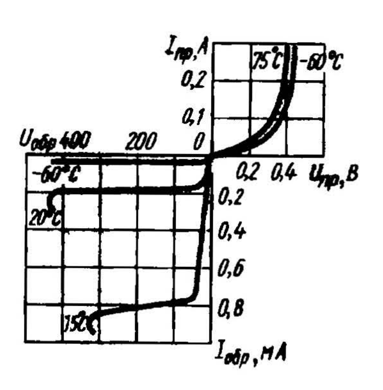


Рис. 3.41. ВАХ одного из германиевых выпрямительных диодов при разных температурах окружающей среды

3. Значительно большая плотность обратного тока в германиевых диодах, так как при прочих равных условиях концентрация неосновных носителей заряда в германии больше на несколько порядков, чем в кремнии.

4. Бóльшие обратные токи через германиевые диоды, в результате чего пробой германиевых диодов имеет тепловой характер. Поэтому пробивное напряжение германиевых диодов уменьшается с увеличением температуры, а значения этого напряжения меньше пробивных напряжений кремниевых диодов.

5. Верхний предел диапазона рабочих температур германиевых диодов составляет приблизительно 75°С, что значительно ниже по сравнению с тем же параметром кремниевых диодов.

Существенной особенностью германиевых диодов и их недостатком является то, что они плохо выдерживают даже очень кратковременные импульсные перегрузки при обратном направлении для *p-n*-перехода. Определяется это механизмом пробоя германиевых диодов — тепловым пробоем, происходящим при шнуровании тока с выделением большой удельной мощности в месте пробоя.

Арсенид-галлиевые диоды

Все перечисленные особенности германиевых и кремниевых выпрямительных диодов в итоге связаны с различием ширины за-

прещенной зоны исходных полупроводниковых материалов. Из приведенного сопоставления видно, что выпрямительные диоды, изготовленные из полупроводникового материала с большой шириной запрещенной зоны, обладают существенными преимуществами в свойствах и параметрах. Одним из таких материалов является арсенид галлия, ширина запрещенной зоны которого при комнатной температуре $\Delta Э = 1{,}43$ эВ.

Параметры первых типов арсенид-галлиевых выпрямительных диодов, выпускаемых промышленностью, еще далеки от оптимально возможных. Так, арсенид-галлиевые диоды типа АД112А рассчитаны на максимально допустимый прямой ток 300 мА при прямом напряжении не более 3 В. Прямое напряжение велико, что является недостатком вообще всех выпрямительных диодов, *p-n*-переходы которых сформированы в материале с большой шириной запрещенной зоны. Максимально допустимое обратное напряжение диодов этой марки — всего 50 В. Низкое значение пробивного напряжения и соответственно максимально допустимого обратного напряжения вызвано, вероятно, большой концентрацией дефектов в области *p-n*-перехода.

Положительными особенностями арсенид-галлиевых выпрямительных диодов являются значительно больший диапазон рабочих температур и лучшие частотные свойства. Верхний предел диапазона рабочих температур арсенид-галлиевых диодов АД112А составляет 250°С. Арсенид-галлиевые диоды АД110А могут работать в качестве выпрямителей малой мощности до частоты 1 МГц, что обеспечивается малым временем жизни носителей заряда в этом материале. Таким образом, арсенид-галлиевые выпрямительные диоды по своим частотным свойствам выходят за пределы низкочастотного диапазона.

§ 3.20. СЕЛЕНОВЫЕ ВЫПРЯМИТЕЛИ

Технология изготовления и конструкция. Селеновые выпрямительные пластины (рис. 3.42) изготовляют на алюминиевых основаниях, которые являются одним из токосъемных электродов. Для уменьшения переходного сопротивления между алюминиевым основанием и наносимым впоследствии слоем селена алюминиевые основания подвергают электрохимическому травлению. Затем на основание наносят аморфный селен. Следующей технологической операцией является термообработка при температуре 215°С, близкой к температуре плавления селена. При этом происходит кристаллизация селена с уменьшением его удельного сопротивления на несколько порядков. Толщина слоя селена составляет 50...60 мкм.

Для создания второго токосъемного электрода на поверхность слоя селена наносят легкоплавкий сплав из висмута, кадмия и олова. При нанесении этого сплава кадмий реагирует с селеном и образуется тонкий слой селенида кадмия. Таким образом, выпрямляющим электрическим переходом в селеновых выпрямителях является гетеропереход между селеном и селенидом кадмия.

Для улучшения выпрямляющих свойств селеновые выпрямительные пластины подвергают электрической формовке, которую производят путем длительного приложения постоянного напряжения в обратном направлении. При этом создаются благоприятные условия для диффузии кадмия в селен и наращивания слоя селенида кадмия. Прохождение электрического тока способствует образованию равномерного по толщине выпрямляющего электрического перехода, так как в наиболее слабых местах перехода будет бóльшая плотность обратного тока, что приведет к локальному повышению температуры и более интенсивному соединению кадмия с селеном.

В установках для выпрямления переменного тока соединяют ряд селеновых выпрямительных пластин последовательно для получения более высокого выпрямленного напряжения и параллельно для выпрямления больших токов. Промышленность выпускает селеновые выпрямительные столбы с последовательным соединением до 1440 селеновых пластин в единой конструкции (выпрямитель 15ГЕ144ОУ-С). К такому выпрямительному столбу может быть приложено переменное напряжение 40 кВ.

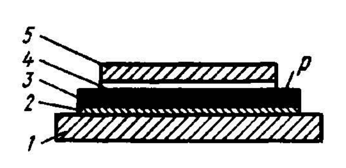


Рис. 3.42. Структура селеновой выпрямительной пластины:
1 — алюминиевое основание; *2* — тонкий слой висмута или никеля, нанесенный на протравленное алюминиевое основание; *3* — слой селена; *4* — слой селенида кадмия; *5* — электрод из сплава висмута, кадмия и олова

Площадь отдельных селеновых выпрямительных пластин, выпускаемых промышленностью, 0,1 ... 400 см^2. Параллельное соединение выпрямительных пластин позволяет получить выпрямленный ток 500 А от одного выпрямительного блока (выпрямитель 140ГЖ24Я4У).

В технологии изготовления селеновых выпрямителей различных серий могут существовать свои особенности. В соответствии с этими особенностями полярность прямого включения может быть противоположной по сравнению с указанной на рис. 3.42, т. е. выпрямляющий электрический переход может быть создан вблизи алюминиевого основания или вблизи верхнего электрода.

Электрические свойства. ВАХ селеновой выпрямительной пластины представлена на рис. 3.43.

Максимально допустимая плотность тока в прямом направлении для селеновых выпрямителей разных серий не превосходит 100 мА/см^2, что на три порядка меньше по сравнению с кремниевыми и германиевыми диодами. Именно поэтому для выпрямления больших токов приходится соединять селеновые выпрямительные пластины параллельно даже при наличии пластин большой площади.

Максимальное прямое напряжение на одной селеновой выпрямительной пластине 0,45...0,75 В (для разных групп). Но в связи с тем что максимально допустимое обратное напряжение селеновых выпрямительных пластин не превосходит 60 В, для выпрям-

ления больших напряжений необходимо соединять последовательно большое число пластин. При этом прямое напряжение на селеновом выпрямительном столбе возрастает во столько раз, сколько пластин соединено последовательно.

Большая барьерная емкость селеновых выпрямителей при относительно малой допустимой плотности прямого тока ограничивает их применение при повышенных частотах.

Параметры селеновых выпрямителей изменяются со временем как при хранении, так и при работе. Длительное хранение приводит к возрастанию обратного тока. Этот процесс, известный под названием расформовки, вызван диффузией галоидов и некоторым изменением химического состава полупроводника. Расформовка, как правило, процесс обратимый. После подачи на выпрямитель обратного или переменного напряжения обратный ток уменьшается со временем, достигая номинального значения в течение 2...3 мин.

Необратимое возрастание прямого напряжения на селеновых выпрямителях, называемое старением, при хранении незначительно, но ускоряется при эксплуатации. Процесс старения связан с увеличением сопротивления слоя селена вследствие обеднения его галогенными примесями, а также с увеличением сопротивления омнческого перехода между селеном и алюминиевым электродом. Интенсивность этих процессов возрастает с повышением температуры, что и определяет верхний предел диапазона рабочих температур селеновых выпрямителей, который составляет для разных серий выпрямителей 75 ... ... 125°C.

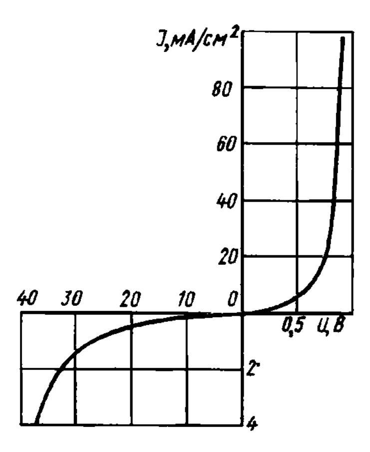


Рис. 3.43. ВАХ селеновой выпрямительной пластины

Таким образом, по многим параметрам и свойствам селеновые поликристаллические выпрямители значительно уступают монокристаллическим кремниевым и германиевым выпрямительным диодам. Однако применение селеновых выпрямителей в различных установках остается широким, а промышленный выпуск различных типов селеновых выпрямителей не снижается. Вызвано это, во-первых, простотой технологии изготовления селеновых выпрямителей и, следовательно, низкой их себестоимостью.

Во-вторых, положительной особенностью селеновых выпрямителей является их способность выдерживать значительные кратковременные перегрузки по току и напряжению и быстро восстанавливать свои свойства после пробоя. Тепловой пробой происходит при шнуровании тока с резким увеличением плотности обратного

тока в месте пробоя. Из-за увеличения удельной выделяющейся мощности в месте пробоя происходит расплавление селена, который затем при быстром охлаждении переходит в аморфное состояние. Удельное сопротивление аморфного селена на несколько порядков выше, чем кристаллического. В результате пробитое место выпрямительной пластины оказывается изолированным, происходит «самозалечивание» места пробоя.

§ 3.21. ИМПУЛЬСНЫЕ ДИОДЫ

Импульсный полупроводниковый диод — это полупроводниковый диод, имеющий малую длительность переходных процессов и предназначенный для применения в импульсных режимах работы.

Основное назначение импульсных диодов — работа в качестве коммутирующих элементов электронных вычислительных машин. Кроме того, импульсные диоды широко применяют в радиоэлектронике для детектирования высокочастотных сигналов и для других целей. Условия работы импульсных диодов обычно соответствуют высокому уровню инжекции, т. е. относительно большим прямым токам. Поэтому свойства и параметры импульсных диодов определяются переходными процессами, рассмотренными в § 3.18.

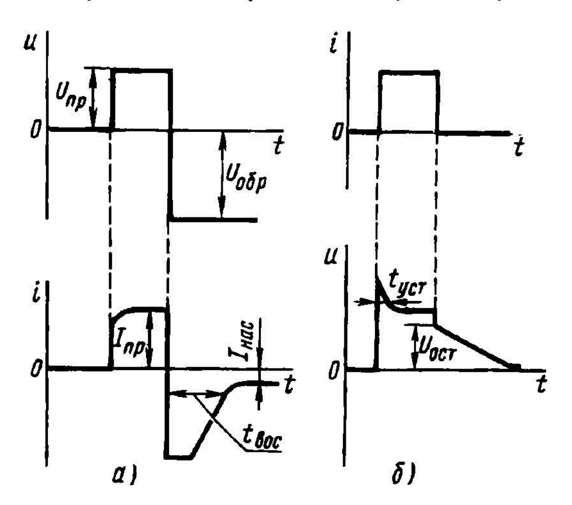


Рис. 3.44. Осциллограммы токов и напряжений импульсного диода при его работе в схемах с генератором напряжения (*а*) и с генератором тока (*б*)

При переключении диода с прямого направления на обратное в начальный момент времени через диод идет большой обратный ток, ограниченный в основном объемным сопротивлением базы (при идеальном генераторе напряжения). С течением времени накопленные в базе неосновные носители заряда рекомбинируют или уходят из базы через *p-n*-переход, после чего обратный ток уменьшается до своего стационарного значения (рис. 3.44, *а*). Переход-

ный процесс, в течение которого обратное сопротивление полупроводникового диода восстаиавливается до постояниого значения после быстрого переключения с прямого иаправления иа обратное, называют восстановлением обратиого сопротивлеиия диода.

Соответственно одним из осиовных параметров импульсиого диода является *время восстановления обратного сопротивления* $t_{\text{вос}}$, равиое интервалу времеии от момента прохождения тока через нуль после переключеиия диода с заданного прямого тока в состояние заданного обратного напряжения до момента достижеиия обратным током заданиого низкого значения. По зиачеиию этого параметра все импульсиые диоды подразделяют на шесть групп, характеризующихся временем восстаиовления обратного сопротивления более 500 ис, 150...500, 30...150, 5...30, 1...5 и менее 1 ис.

При пропускании импульса тока в прямом иаправлении через диод иаблюдается выброс напряжеиия в первый момент после включеиия (рис. 3.44, *б*). Вызвано это повышеиым падеиием напряжеиия до тех пор, пока ие произойдет накопление неосиовиых иосителей заряда в базе диода в результате инжекции и не уменьшится объемиое сопротивление базы. Переходный процесс, в течение которого прямое сопротивлеиие полупроводникового диода устаиавливается до постояиного зиачеиия после быстрого включеиня диода в прямом иаправлении, называют установлеиием прямого сопротивлеиия диода.

Соответственно следующим специфнческим параметром импульсного днода является *время установления прямого напряжения диода* $t_{\text{уст}}$, равное интервалу времени от момеита подачи импульса прямого тока на диод (при нулевом начальиом напряжении смещеиия) до достижения заданного значения прямого напряжеиия на диоде.

Значения этих параметров ($t_{\text{вос}}$ и $t_{\text{уст}}$) зависят от структуры диода, от времени жизни неосиовных носителей заряда в базе диода, а также от условий измерения. Так, прн увеличеини прямого тока, проходящего через диод до переключения иа обратное напряжение, время восстановления обратного сопротивления увеличивается, что вызвано иеобходимостью рассасывания большего количества накопленных в базе неосновных носителей заряда. При увеличенин обратиого иапряжения по абсолютиому значеиию время восстаиовлеиия обратного сопротивления уменьшается, т. е. обратное напряжеиие способствует процессу рассасывания неосиовиых иосителей заряда из базы диода. Одиако если выброс обратного тока при переключении днода вызваи перезарядкой барьерной емкости (см. § 3.18), то время перезарядки возрастает с увеличением обратного напряжения, что соответствует увеличеиию времени восстаиовлення обратного сопротивлеиия диода.

Одной из первых была разработаиа конструкция точечного импульсного диода (рис. 3.45). Точечный диод состоит из кристалла германия, припаяниого к кристаллодержателю, контактного электрода в вцде тонкой проволокн и стекляииого баллоиа.

Точечиый *p-n*-переход в таких диодах может быть сформиро-

ван с помощью импульса тока, при прохождении которого выделяющаяся мощность разогревает область полупроводника под точечным контактом, а тип электропроводности этой проконтактной области изменяется из-за внедрения примесных ионов. Такой метод (метод электроформовки) дает возможность получить полусферический *p-n*-переход с радиусом около 20 мкм и с малым временем жизни неосновных носителей заряда в базе под переходом. Другой метод отличается тем, что при прохождении импульса тока происходит приварка проволоки к полупроводнику с образованием аналогичного *p-n*-перехода.

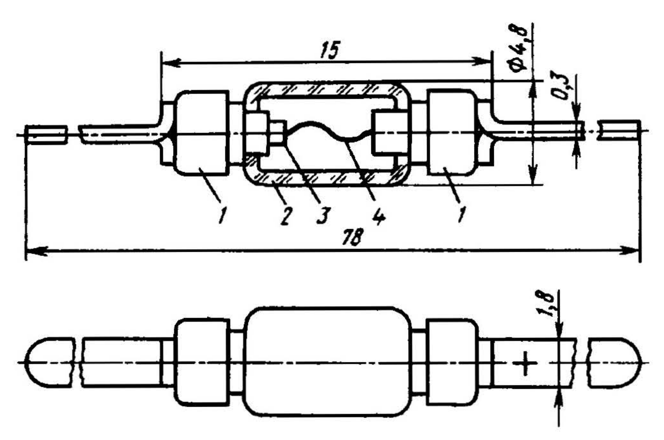


Рис. 3.45. Конструкция одного из импульсных точечных диодов:

1 — выводы; *2* — стеклянный баллон; *3* — кристалл германия; *4* — электрод из тонкой проволоки

В связи с малым временем жизни неосновных носителей заряда время восстановления обратного сопротивления точечных диодов значительно меньше, чем плоскостных выпрямительных диодов. Барьерная емкость точечного *p-n*-перехода мала из-за малой площади этого перехода. Поэтому точечные диоды сохраняют выпрямительные свойства до десятков мегагерц.

Особенностью точечных диодов является большое сопротивление базы, которое определяется в этих диодах *сопротивлением растекания* r_s, т. е. сопротивлением области полупроводника, расположенной вблизи точечного *p-n*-перехода, где сгущаются линии тока. Так как сопротивление растекания необходимо учитывать и в других структурах различных полупроводниковых приборов, то представляет интерес вывод общей формулы для этого сопротивления.

Определим сопротивление кристалла полупроводника под полусферическим *p-n*-переходом или под любым другим полусферическим электрическим переходом с радиусом a (рис. 3.46). Сопротивление полусферического слоя полупроводника толщиной dR на расстоянии R от центра сферы

$$dr = \varrho \frac{dR}{2\pi R^2}.$$

Полное сопротивление кристалла полупроводника толщиной b под точечным переходом есть сумма сопротивлений полусферических слоев. Если удельное сопротивление ϱ считать постоянным, то

$$r_s = \int_0^{r_s} dr = \frac{\varrho}{2\pi}\int_a^b \frac{dR}{R^2} = \frac{\varrho}{2\pi}\left(\frac{1}{a} - \frac{1}{b}\right).$$

Если радиус a значительно меньше толщины кристалла b (в точечных диодах $a = 5...20$ мкм, $b \approx 200$ мкм), то, пренебрегая значением $1/b$, получим

$$r_s \approx \varrho/(2\pi a). \qquad (3.103)$$

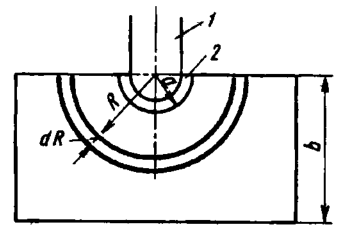


Рис. 3.46. Структура точечного диода:

1 — проволочный электрод; 2 — слой полупроводника с электропроводностью p-типа

Для получения достаточно больших значений пробивного напряжения p-n-перехода исходный полупроводниковый материал должен иметь большое удельное сопротивление. Но при этом будет велико и сопротивление растекания (сопротивление базы точечного днода), что приведет к увеличению прямого напряжения на диоде.

В связи с существенными недостатками точечных диодов они практически полностью вытесиены импульсными диодами, производство которых основано на современных производительных и контролируемых методах формироваиия p-n-переходов с использованием планарной технологии, эпитаксиальиого наращиваиия, а также ионно-лучевой технологии. Основным исходным полупроводниковым материалом при этом служит кремний, а иногда и арсеиид галлия.

Для ускорения переходных процессов в кремниевых импульсных диодах и для уменьшеиия значения времени восстановления обратного сопротивления этих диодов в исходный кремний вводят примесь золота. Эта примесь обеспечивает оявление в запрещеиной зоне кремния энергетических уровней рекомбинационных ловушек и уменьшение времени жизни неосновных носителей.

При изготовлении планарных импульсных диодов диффузия примесей проводится через отверстия или окна в слое оксида кремния (см. § 2.4). Площадь p-n-перехода таких диодов может быть сделана достаточно малой, мала будет и барьериая емкость.

Планарная технология дает возможность относительно просто сформировать на одном кристалле много диодных структур. Таким образом делают *наборы* (сборки или матрицы) *импульсных диодов*, т. е. совокупность импульсных диодов, собранных в единую

конструкцию, не соединенных электрически или соединенных по одноименным выводам. Наборы импульсных диодов удобны для применения в установках вычислительной техники, отдельные узлы которых содержат большое число однотипных импульсных диодов.

§ 3.22. ДИОДЫ ШОТКИ

Диод Шотки — это полупроводниковый диод, выпрямительные свойства которого основаны на использовании выпрямляющего электрического перехода между металлом и полупроводником.

Для всех диодов, рассмотренных в предыдущих параграфах, основным физическим процессом, ограничивающим диапазон рабочих частот, оказывался процесс накопления и рассасывания неосновных носителей заряда в базе диода. Другой физический процесс — перезаряд барьерной емкости выпрямляющего электрического перехода — имел в рассмотренных диодах второстепенное значение и сказывался на их частотных свойствах только при определенных условиях. Поэтому были выдвинуты требования к конструкции и технологии изготовления диодов, выполнение которых обеспечивало бы ускорение рассасывания накопленных в базе за время действия прямого напряжения неосновных носителей заряда. Понятно, что если исключить инжекцию неосновных носителей заряда при работе диода, то не было бы накопления этих неосновных носителей в базе и соответственно относительно медленного процесса их рассасывания. Здесь можно перечислить несколько возможностей практически полного устранения инжекции неосновных носителей заряда при сохранении выпрямительных свойств полупроводниковых диодов.

1. Использование в качестве выпрямляющего электрического перехода гетерогенного перехода (гетероперехода), т. е. электрического перехода, образованного в результате контакта полупроводников с различной шириной запрещенной зоны. Инжекция неосновных носителей при прямом включении будет отсутствовать при выполнении ряда условий и, в частности, при одинаковом типе электропроводности полупроводников, образующих гетеропереход (см. § 2.11). Этот способ устранения инжекции неосновных носителей заряда пока не нашел широкого применения в промышленном производстве монокристаллических полупроводниковых диодов из-за технологических трудностей.

2. Использование для выпрямления эффекта туннелирования (см. § 3.30).

3. Инвертирование диодов, т. е. использование для выпрямления только обратной ветви ВАХ вместе с участком, соответствующим лавинному пробою. Этот способ не нашел применения из-за необходимости иметь для каждого диода свое напряжение смещения, почти равное напряжению пробоя. Кроме того, в начальной стадии лавинного пробоя в диоде возникают шумы (см. § 3.25).

4. Использование выпрямляющего перехода Шотки, т. е. выпрямляющего электрического перехода, образованного в результа-

те контакта между металлом и полупроводником. На таком переходе высота потенциального барьера для электронов и дырок может существенно отличаться, как показано в § 2.10. Поэтому при включении выпрямляющего перехода Шотки в прямом направлении прямой ток возникает благодаря движению основных носителей заряда полупроводника в металл, а носители другого знака (неосновные для полупроводника) практически не могут перейти из металла в полупроводник из-за высокого для них потенциального барьера на переходе (см. рис. 2.16).

Таким образом, на основе выпрямляющего перехода Шоткн могут быть созданы выпрямительные, нмпульсные и сверхвысокочастотные полупроводниковые диоды, отличающиеся от диодов с p-n-переходом лучшнми частотными свойствами.

Выпрямительные диоды Шотки. На частотные свойства диодов Шотки основное влияние должно оказывать время перезарядки барьерной емкости перехода. Постоянная времени перезарядки за-

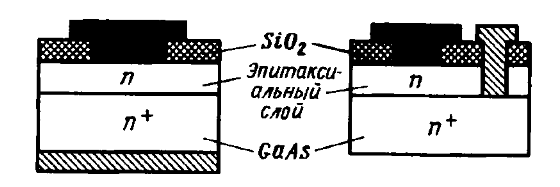


Рис. 3.47. Варианты структур диодов Шотки с двухслойной базой

висит и от сопротивления базы диода ($\tau = r_б C_{бар}$). Поэтому выпрямляющий переход Шотки целесообразиее создавать на кристалле полупроводника с электропроводностью n-типа — подвижность электронов больше подвижности дырок. По той же причине должна быть большой и концентрация примесей в кристалле полупроводника.

Однако толщина потенциального барьера Шотки, возникающего в полупроводнике вблизи границы раздела с металлом, должна быть достаточно большой. Только при большой толщине потенциального барьера (перехода Шотки) можно будет, во-первых, устранить вероятность туннелирования носнтелей заряда сквозь потенциальный барьер, во-вторых, получить достаточные значения пробивного напряжения и, в-третьих, получить меньшие значения удельиой (на единицу площади) барьерной емкости перехода. А толщина перехода или потенциального барьера зависит от концентрации примесей в полупроводнике: чем больше концентрация примесей, тем тоньше переход (см. § 2.6). Отсюда следует противоположное требование меньшей концентрации примесей в полупроводнике.

Учет этих противоречивых требований к концентрации примесей в исходном полупроводнике приводит к необходимости созда-

ния двухслойной базы диода Шотки (рис. 3.47). Основная часть кристалла — подложка толщиной около 0,2 мм — содержит большую концентрацию примесей и имеет малое удельное сопротивление. Тонкий монокристаллический слой того же самого полупроводинка (толщиной в несколько микрометров) с той же электропроводностью n-типа может быть получен на поверхности подложки методом эпитаксиального наращивания. Концентрация доноров в эпитаксиальном слое должна быть значительно меньше, чем концентрация доноров в подложке.

В качестве исходного полупроводникового материала для выпрямительных диодов Шотки можно использовать кремний или арсенид галлия. Однако в эпитаксиальных слоях арсенида галлия не удается пока достичь малой концентрации дефектов и достаточно низкой концентрации доноров. Поэтому пробивное напряжение диодов Шотки на основе арсенида галлия оказывается низким, что является существенным недостатком для выпрямительных диодов.

Металлический электрод на эпитаксиальный слой полупроводника обычно наносят методом испарения в вакууме с последующим осаждением на поверхность эпитаксиального слоя. Перед нанесением металлического электрода целесообразно методами фотолитографии создать окна в оксидном слое на поверхности полупроводника. Так легче получить выпрямляющий переход Шотки необходимых площади и конфигурации.

Выпрямительные низкочастотные диоды предпочтительнее изготовлять с p-n-переходом. Выпрямительные диоды Шотки в области низких частот могут в перспективе иметь преимущество перед диодами с p-n-переходом, связанное с простотой изготовления.

Наибольшие преимущества перед диодами с p-n-переходом диоды Шотки должны иметь при выпрямлении больших токов высокой частоты. Здесь кроме лучших частотных свойств диодов Шотки следует отметить такие их особенности: меньшее прямое напряжение из-за меньшей высоты потенциального барьера для основных носителей заряда полупроводника; бóльшая максимально допустимая плотность прямого тока, что связано, во-первых, с меньшим прямым напряжением и, во-вторых, с хорошим теплоотводом от выпрямляющего перехода Шотки. Действительно, металлический слой, находящийся с одной стороны перехода Шотки, по своей теплопроводности превосходит любой сильнолегированный слой полупроводника. По этим же причинам выпрямительные диоды Шотки должны выдерживать значительно большие перегрузки по току по сравнению с аналогичными диодами с p-n-переходом на основе того же самого полупроводникового материала.

Еще одна особенность диодов Шотки заключается в идеальности прямой ветви ВАХ — прямая ветвь ВАХ соответствует выражению (3.27). При этом с изменением прямого тока в пределах нескольких порядков зависимость $\lg I_{пр} = f(U_{пр})$ близка к линей-

ной, или в показателе экспоненты при изменении тока не появляется дополнительных множителей (см. § 3.17 и рис. 3.32). Учитывая эту особенность, диоды Шотки можно использовать в качестве быстродействующих логарифмических элементов.

На рис. 3.48 показаны ВАХ кремниевого диода Шотки 2Д219, рассчитанного на максимально допустимый прямой ток 10 А. Прямое напряжение на диоде при максимально допустимом прямом токе не более 0,6 В, максимально допустимое обратное напряжение для диода 2Д219Б 20 В. Эти диоды допускают прохождение импульсов тока длительностью до 10 мс с периодом повторения не менее 10 мин с амплитудой, в 25 раз превышающей максимально допустимый прямой ток. Диоды рассчитаны на частоту выпрямляемого тока 0,2 МГц.

Импульсные диоды Шотки. Исходным полупроводниковым материалом для этих диодов может быть, так же как и для выпрямительных диодов Шотки, кремний или арсенид галлия. Но предпочтение здесь должно быть отдано арсениду галлия, так как в этом материале время жизни неосновных носителей заряда может быть менее 10^{-9} с. Несмотря на практическое отсутствие инжекции неосновных носителей заряда через переход Шотки при его включении в прямом направлении (что уже было отмечено ранее), при больших прямых напряжениях и плотностях прямого тока существует, конечно, некоторая составляющая прямого тока, связанная с инжекцией неосновных носителей заряда в полупроводник. Поэтому требование малости времени жизни неосновных носителей в исходном полупроводниковом материале остается и для импульсных диодов Шотки.

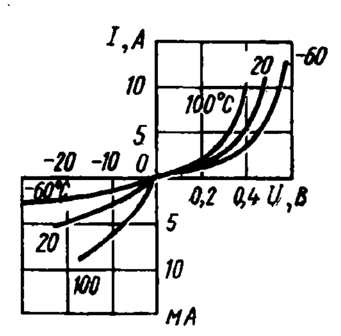


Рис. 3.48. ВАХ кремниевого диода Шотки 2Д219 при разных температурах

Арсенид галлия пока не удается получить с малой концентрацией дефектов, в результате чего арсенид-галлиевые диоды имеют относительно малые значения пробивных напряжений, далекие от теоретически возможных. Это является существенным недостатком для выпрямительных диодов, но не столь важно для импульсных диодов, так как большая часть импульсных схем — это низковольтные схемы.

Выпускаемые промышленностью арсенид-галлиевые импульсные диоды Шотки (3А527А, 3А530Б и др.) предназначены для использования в импульсных схемах пико- и наносекундного диапазона. В отличие от выпрямительных диодов Шотки они имеют значительно меньшие площади выпрямляющих переходов. Поэтому общая емкость этих диодов не превышает 1 пФ даже при нулевом постоянном напряжении смещения.

Сверхвысокочастотные диоды Шотки будут рассмотрены в § 3.24, хотя арсенид-галлиевые импульсные диоды перечисленных марок по их частотным свойствам тоже можно считать СВЧ-диодами, о чем свидетельствует и их маркировка.

§ 3.23. ДИОДЫ С РЕЗКИМ ВОССТАНОВЛЕНИЕМ ОБРАТНОГО СОПРОТИВЛЕНИЯ

Диод с резким восстаиовлеиием обратиого сопротивлеиия — это полупроводииковый диод, у которого используется эффект резкого восстаиовления обратного сопротивлеиия для целей умиожеиия частоты и формирования импульсов с малым времеием иарастаиия.

Как видно из предыдущих параграфов (см. § 3.18, 3.19, 3.21 и 3.22), при конструировании быстродействующих выпрямительных и импульсных диодов основные усилия должны быть направлены на сокращенне длительности переходных процессов и прежде всего на ускорение процесса рассасывания или на практическое исключение эффекта накопления иеосиовиых носителей заряда в базе путем исключеиия инжекции неосновных носителей.

Однако, как это часто бывает, нежелательный для одиих приборов физический процесс удается использовать в качестве основы прииципа действия других приборов. Так произошло и с процессом рассасывания накопленных в базе диода неосновных носителей заряда.

Если при прямом включении происходила иижекция иеосновных носителей заряда в базу диода, то при его переключеиии на обратное напряжение будет происходить рассасывание накопленных неосновных иосителей. В первое время после переключеиия сопротивлеиие диода определяется только объемным сопротивлением базы. Процесс рассасывания неосновных носителей заряда можно разбить на две фазы (рис. 3.49).

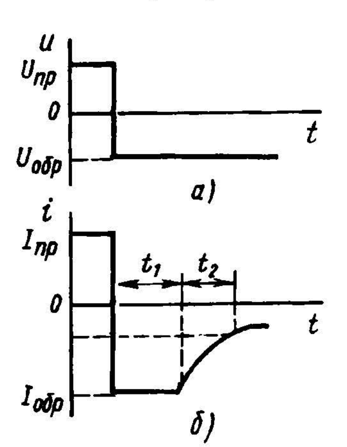


Рис. 3.49. Осциллограммы напряжеиия (*а*) и тока (*б*) при переключеиии диода с прямого иапряжения на обратное:
t_1 — фаза высокой обратной проводимости; t_2 — фаза спада обратного тока

Длительность первой фазы t_1 — фазы высокой обратной проводимости — определяется временем от момента прохождения тока через нуль при переключении диода до начала спада обратного тока. За время первой фазы в базе диода около *p-n*-перехода происходит умеиьшение граничной концентрацин неосновных носителей заряда до иуля (см. рис. 3.33, *е*). Длительность этой фазы за-

висит от количества накопленных в базе неосновных носителей, т. е. от прямого тока, предшествующего переключению, и от амплитуды обратного тока. Амплитуда обратного тока, в свою очередь, зависит от обратного напряжения, поданного на диод при переключении, и от сопротивления базы при условии, что диод работает в схеме с идеальным генератором напряжения. В реальных условиях амплитуда обратного тока определяется ЭДС генератора обратного напряжения и суммой сопротивлений базы диода и внешней цепи.

Длительность второй фазы t_2 — фазы спада обратного тока — определяется временем уменьшения обратного тока до определенного малого значения. За время второй фазы происходит дальнейшее рассасывание неосновных носителей заряда из глубинных слоев базы, а также рекомбинация их в базе.

При малой длительности второй фазы импульс обратного тока, полученный при переключении диода, имеет почти прямоугольную форму. Использование диодов с резким восстановлением обратного сопротивления в схемах формирователей коротких импульсов большой амплитуды — одно из основных назначений этих диодов. Другая возможность применения связана с тем, что короткий импульс большой амплитуды содержит много гармоник высшего порядка, которые могут быть выделены схемным путем. Таким образом можно осуществить умножение частоты электромагнитных колебаний.

Для формирования импульсов и умножения частоты можно использовать и другие импульсные диоды, но они должны быть с инжекцией неосновных носителей заряда и с малой длительностью спада обратного тока, т. е. к структуре и конструкции диодов с резким восстановлением обратного сопротивления предъявляются свои специфические требования.

В диоде, изготовленном методом диффузии примеси, неравномерное распределение нескомпенсированной примеси может существовать и за пределами p-n-перехода в базе. Для базы с электропроводностью n-типа при отсутствии тока

$$J_n = qn\mu_n E + qD_n \operatorname{grad} n = 0.$$

Отсюда электрическое поле в базе за пределами p-n-перехода

$$E = -\frac{D_n}{\mu_n}\frac{\operatorname{grad} n}{n}.$$

Учитывая полную ионизацию доноров и используя соотношение Эйнштейна, получим

$$E = -\frac{kT}{q}\frac{\operatorname{grad} N_д}{N_д}. \tag{3.104}$$

Возникновение электрического поля в полупроводнике при наличии градиента концентрации примесей вызвано диффузией носителей заряда (для полупроводника с электропроводно-

стью n-типа — электронов) в места с меньшей концентрацией. В результате диффузии носителей заряда нарушается электрическая нейтральность различных частей полупроводника: с одной стороны, получается избыток носителей заряда (в нашем примере — электронов), с другой стороны, остаются ионизированные нескомпенсированные примеси (в нашем примере — доноры, т. е. положительные неподвижные заряды). После возникновения между разноименными зарядами электрического поля, которое в данном случае часто называют встроеним полем, устанавливается динамическое равновесие между диффузией и дрейфом носителей заряда.

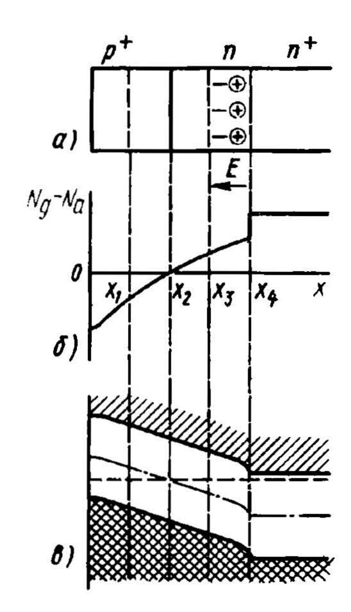


Рис. 3.50. Структура эпитаксиального диода *(а)*, распределение примесей в структуре *(б)* и энергетическая диаграмма этой структуры *(в)*: x_1 — граница p-n-перехода в p-области; x_2 — граница между p- и n-областями (металлургическая граница); x_3 — граница p-n-перехода в n-области; x_4 — толщина эпитаксиального слоя; E — напряженность встроенного электрического поля в высокоомной части базы

Существование электрнческого поля при наличии градиента концентрации примесей можно также проиллюстрировать с помощью энергетических диаграмм. Прн термоднамнческом равновесии уровень Ферми должен быть горизонтальным. Разрешенные энергетические зоны должны быть расположены относительно уровня Ферми на расстояини в единицах энергии, зависящем от концентрацин примесей. При неравномерном распределении примесей края энергетических зон (дно зоны проводимости и потолок валентной зоны) получаются наклонными, что соответствует наличию электрического поля. Энергетическая диаграмма диодной структуры с p-n-переходом, сформированным методом диффузии акцепторов в высокоомном эпитаксиальном слое, который выращен на низкоомной подложке, показана на рис. 3.50.

Таким образом, в базе диода около p-n-перехода может существовать электрическое поле, которое обычно называют встроенным, так как оно возникает в процессе диффузии примесей и существует при отсутствии внешних напряжений, т. е. прн отсутствии тока. В диодной структуре, показанной на рис. 3.50, встроенное поле тормозит неосновные носители заряда — дырки, инжектнрованные через переход при прямом включенин, и ускоряет их движение к переходу во время рассасывання при обратном включении. При этом накопленне неосновных носителей должно происходить только вблизи p-n-перехода, а длительность спада обратного тока должна быть малой.

Для получения существенного ускореиия процесса рассасывания неосновных носителей заряда необходима большая напряженность встроенного поля н в соответствии с (3.104) большой градиент концентрации примесей. Однако увеличение градиента концентрации примесей приводит к уменьшению пробивного напряжения (см. § 3.11) и увелнчению барьерной емкости p-n-перехода (см. § 2.8).

Кроме того, при высоком уровне инжекцни в базе диода около p-n-перехода возникает электрическое поле, пропорциональное градиенту концентрации инжектированных неосновных иосителей заряда (см. § 3.15). Эта составляющая электрического поля направлена навстречу встроенному полю, а ее значение может значительно превышать значение встроенного поля.

Для получения минимального сопротивления базы подложку низкоомного кремния выбирают с большой концентрацией доноров. Между низкоомной подложкой и высокоомным эпитаксиальным слоем, выращенным на подложке, возникает потенциальный барьер для неосновных носителей — дырок, инжектированных через p-n-переход при его прямом включенни. Такнм образом, омический переход между подложкой и эпитаксиальным слоем (рис. 3.50) также выполняет роль ограничителя области накоплення неосновных носителей заряда в базе диода с резким восстановлением обратного сопротнвления.

Наибольший практический интерес представляют диоды с резким восстановлением. обратного сопротнвления, пригодные для формирования прямоугольных импульсов тока с нано- и пикосекундной длительностью, так как именно в этом диапазоне диоды с резким восстановлением обратного сопротивления позволяют получать не достижимые другнми способами результаты. При конструировании таких диодов, предназначенных для СВЧ-диапазона, необходимо предусмотреть возможное уменьшение емкости и индуктивностн элементов корпуса, внутренних и внешних выводов.

Кроме параметров, применяемых и для других полупроводниковых диодов, для оценки качества и функциональных возможностей диодов с резким восстановлением обратного сопротивления используют следующие специфические параметры:

1. *Эффективное время жизни неосновных носителей заряда* $t_{эф}$ — время, определяющее процесс рекомбинацни неосновных носителей заряда в базе диода.

2. *Заряд переключения* $Q_{пк}$ — часть накопленного заряда, вытекающая во внешнюю цепь при изменении направления тока с прямого на обратное. Заряд переключения зависит от прямого тока, предшествующего переключению диода, и от обратного напряжения, приложенного к диоду после переключения. Поэтому при указании значения заряда переключения должны быть указаны и условия его измерения.

3. *Максимально допустимый импульсный обратный ток* $I_{обр\,и\,max}$ — значение импульсного обратного тока, которое обеспе-

чивает необходимую надежность работы диода. Необходимость введения этого параметра диктуется спецификой работы диодов с резким восстановлением обратного сопротивления. Например, для диода с резким восстановлением обратного сопротивления 2Д524Б максимально допустимый импульсный обратный ток при скважности, равной или более 10, составляет 1000 мА.

§ 3.24. СВЕРХВЫСОКОЧАСТОТНЫЕ ДИОДЫ

Сверхвысокочастотный полупроводниковый диод (СВЧ-диод) — это полупроводниковый диод, предназначенный для преобразования и обработки сверхвысокочастотного сигнала.

Полупроводниковые СВЧ-диоды уже длительное время применяют в различной радиоэлектронной аппаратуре и измерительной технике СВЧ-диапазона, т. е. на частотах более 300 МГц. Вначале СВЧ-диоды использовали для детектирования и смешения сигналов. Для этих целей применяли точечные диоды, выпрямляющий электрический переход в которых возникал между кристаллом полупроводника и прижимным металлическим электродом в виде заостренной пружинки. Созданные в последнее время новые типы СВЧ-диодов практически целиком заменили точечные детекторные и смесительные диоды. Они дают возможность решать задачи генерации и усиления электромагнитных колебаний СВЧ-диапазона, умножения частоты, модуляции, регулирования, ограничения сигналов и т. п.

Здесь рассмотрены не все СВЧ-диоды, так как некоторые из рассмотренных в предыдущих параграфах диоды тоже могут работать на сверхвысоких частотах (импульсные диоды, диоды с резким восстановлением обратного сопротивления).

СВЧ-диоды, имеющие специфический принцип действия (туннельные и обращенные диоды, варикапы, лавинно-пролетные диоды, генераторы Ганна), рассмотрены далее.

Смесительные диоды

Смесительный полупроводниковый диод — это полупроводниковый диод, предназначенный для преобразования высокочастотных сигналов в сигнал промежуточной частоты

К смесительному диоду подводится сигнал и напряжение от специального генератора — гетеродина. В связи с нелинейностью ВАХ диода происходит образование сигнала разностной (промежуточной) частоты. Дальнейшее усиление входного сигнала осуществляется на этой промежуточной частоте, которая должна быть выше частот, соответствующих низкочастотным шумам, обратно пропорциональным частоте.

Основным параметром смесительных диодов, определяющим эффективность преобразования входных сигналов высокой ча-

стоты в сигналы промежуточной частоты, является параметр $L_{прб}$, называемый *потери преобразования смесительного диода* и равный отношению мощности СВЧ-сигнала на входе диодной камеры к мощности сигнала промежуточной частоты, выделяемой в нагрузке смесительного диода в рабочем режиме:

$$L_{прб}\,[дБ] = 10\lg\frac{P_{СВЧ}}{P_{пр.ч}}.$$

В большинстве приемных устройств СВЧ-диапазона отсутствуют усилители перед смесителем. Поэтому чувствительность всего приемного устройства, возможность различить полезный сигнал на фоне шумов зависят от уровня шумов смесительного диода. Уровень шумов смесительного диода (и других приборов) оценивают *шумовым отношением* $n_ш$ — отношением номинальной мощности шумов диода в рабочем режиме к номинальной мощности тепловых шумов соответствующего активного сопротивления при той же температуре и одинаковой полосе частот.

Другим параметром, характеризующим шумы смесительного диода и других приборов и систем, является *коэффициент шума* — отношение мощности шумов на выходе к той ее части, которая вызвана тепловыми шумами источника сигнала:

$$F = (P_{сигн}/P_ш)_{вх}/(P_{сигн}/P_ш)_{вых} = \frac{P_{ш\,вых}}{P_{ш.вх}(P_{сигн.вых}/P_{сигн.вх})}.$$

Обобщенным параметром приемного устройства, в смесителе которого использован диод с определенными потерями преобразования и шумовым отношением, является *нормированный коэффициент шума* — значение коэффициента шума приемного устройства со смесительным диодом на входе при коэффициенте шума усилителя промежуточной частоты $F_{упч}$, равном 1,5 дБ:

$$F_{норм\,[дБ]} = L_{прб\,[дБ]} + 10\lg(n_ш + F_{упч} - 1).$$

Одним из вспомогательных параметров смесительных диодов служит *выпрямленный ток* $I_{вп}$ — постоянная составляющая тока, протекающая в выходной цепи диода в рабочем режиме. Этот параметр используется для контроля исправности смесительного диода и гетеродина приемника, от которого на смесительный диод подается определенная мощность СВЧ-колебаний с определенной длиной волны.

Другим вспомогательным параметром является *коэффициент стоячей волны по напряжению* СВЧ-диода $K_{ст\,U}$ — коэффициент стоячей волны по напряжению в передающей линии СВЧ, когда она нагружена на определенную диодную камеру с СВЧ-диодом в рабочем режиме. Чем лучше согласовано входное сопротивление камеры (с диодом) с волновым сопротивлением тракта, тем меньше коэффициент стоячей волны по напряжению и потери принимаемого сигнала.

Основные требования к исходному полупроводниковому мате-

риалу, структуре СВЧ-диодов и причины этих требований можно сформулировать следующим образом.

Условиями получения необходимых частотных свойств СВЧ-диодов являются отсутствие инжекции неосновных носителей заряда через выпрямляющий электрический переход в базу, малое время жизни неосновных носителей в базе и малое значение постоянной времени перезаряда барьерной емкости, т. е. малые барьерная емкость выпрямляющего перехода и сопротивление базы, которое должно быть малым и для уменьшения потерь мощности в диоде.

Пробивное напряжение, хотя оно и не является параметром СВЧ-диодов, должно быть большое. Во-первых, это необходимо для предотвращения детектирования на обратной ветви ВАХ тока от гетеродина, который выдает переменное напряжение с относительно большой амплитудой. Во-вторых, это необходимо для повышения надежности СВЧ-диода, так как на вход радиоприемных устройств могут попадать посторонние радиоимпульсы большой мощности. Для обеспечения достаточных значений пробивного напряжения, а также для уменьшения барьерной емкости концентрация примеси в базе около выпрямляющего электрического перехода должна быть малой, что противоречит требованию малости сопротивления базы.

Для повышения надежности СВЧ-диодов пробой их должен быть лавинным, а не тепловым, который происходит с неизбежным шнурованием тока (см. § 3.13).

Отсюда следует, что исходный полупроводниковый материал для СВЧ-диодов должен иметь большую ширину запрещенной зоны, малое время жизни неосновных носителей заряда и большую подвижность основных носителей заряда, т. е. при заданной концентрации примеси он должен иметь меньшее удельное сопротивление. Таким материалом, в частности, является арсенид галлия.

В качестве смесительных СВЧ-диодов наиболее широкое распространение в последнее время получили диоды Шотки (см. § 3.22). Выпрямляющий электрический переход этих диодов формируют путем напыления металла на поверхность высокоомного эпитаксиального слоя арсенида галлия с использованием фотолитографии.

Размеры выпрямляющего перехода зависят от частоты, на которой должен работать диод. Для очень высоких частот (в десятки и сотни гигагерц, что соответствует миллиметровому диапазону длин волн) с целью уменьшения барьерной емкости необходимо уменьшение диаметра перехода Шотки до 2...3 мкм. Такие размеры оказываются предельными для обычной фотолитографии, в частности, из-за подтравливания защитного слоя диоксида под маской фоторезиста при химическом травлении окон в слое диоксида. Для устранения этого явления используют методы ионно-плазменного травления.

Кроме того, при очень высоких частотах сказывается скин-

эффект в металлическом контакте, образующем переход Шотки. Поэтому с целью увеличения отношения периметра перехода к его площади необходимо создавать кольцевые, полосковые, крестообразные или эллиптические переходы. Для сохранения малой площади перехода, имеющего сложную конфигурацию, ширина кольца, полосок и т. д. должна быть мала (около 1 мкм). Метод фотолитографии при этом оказывается неприемлемым. Здесь используют методы рентгенолитографии и электронно-лучевые методы литографии, которые имеют значительно большую разрешающую способность.

Для удобства включения в соответствующие элементы и цепи СВЧ (например, волноводные и коаксиальные линии передач) СВЧ-диоды оформляют в корпусах различной конструкции (рис. 3.51).

Типы корпусов СВЧ-диодов, их габаритные и присоединительные размеры (как и многих других полупроводниковых приборов) соответствуют ГОСТ 18472—82 (СТ СЭВ 1818—79).

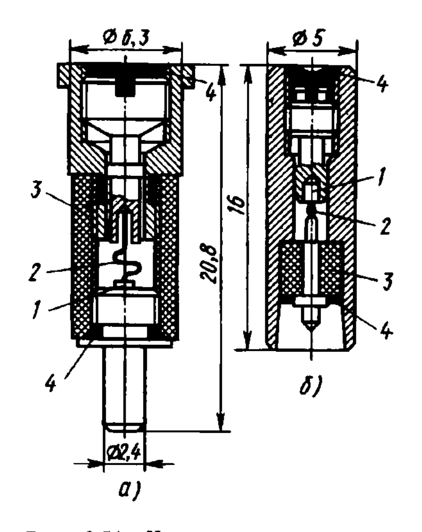


Рис. 3.51. Конструкции некоторых СВЧ-диодов патронного (*а*) и коаксиального типов (*б*):

1 — кристалл полупроводника; *2* — контактная пружина; *3* — керамическая втулка; *4* — герметизирующая заливка

Диоды с патронной конструкцией корпуса, состоящего из керамической втулки и латунных фланцев или ниппелей (рис. 3.51, *а*), предназначены для использования в деци- и сантиметровом диапазонах длин волн, т. е. до частот около 12 ГГц. Диоды с коаксиальной конструкцией корпуса (рис. 3.51, *б*) применяют в диапазоне коротких сантиметровых волн до частот около 30 ГГц. В миллиметровом диапазоне волн используют преимущественно диоды волноводной конструкции, т. е. волноводные вставки, которые и являются своеобразными корпусами СВЧ-диодов. Для полосковых линий и для интегральных СВЧ-микросхем используют либо диоды в миниатюрных корпусах, либо бескорпусные СВЧ-диоды, поверхность полупроводникового кристалла которых защищена только пленкой диоксида.

Конструкция корпуса СВЧ-диода может существенно влиять на его частотные свойства. Для уменьшения этого влияния электрическая емкость корпуса и индуктивность внешних и внутренних выводов должны быть минимальными. Металлические детали СВЧ-диодов обычно покрывают тонким слоем серебра или золота, обеспечивая тем самым минимальное их сопротивление, надежность контакта с внешними цепями и защиту от коррозии.

Детекторные диоды

Детекторный полупроводниковый диод — это полупроводниковый диод, предназначенный для детектирования сигнала.

При детектировании используется выпрямляющее свойство диода для выделения из модулированных по амплитуде ВЧ- или СВЧ-колебаний сигнала более низкой частоты, который потом поступает на вход усилителя (рис. 3.52).

Одним из основных параметров детекторных СВЧ-диодов является *чувствительность по току* β_I — отношение приращения выпрямленного тока при заданной нагрузке в выходной цепи диода к мощности СВЧ-сигнала, подводимой ко входу диодной камеры с детекторным диодом в рабочем режиме и вызвавшей это приращение. Чувствительность по току детекторного диода зависит от постоянного прямого тока смещения (на рис. 3.52 приведено графическое пояснение детектирования без постоянного смещения). Наибольшие значения чувствительности по току обычно получаются при прямом токе смещения в несколько десятков микроампер, но при выборе тока смещения необходимо учитывать его влияние и на другие параметры.

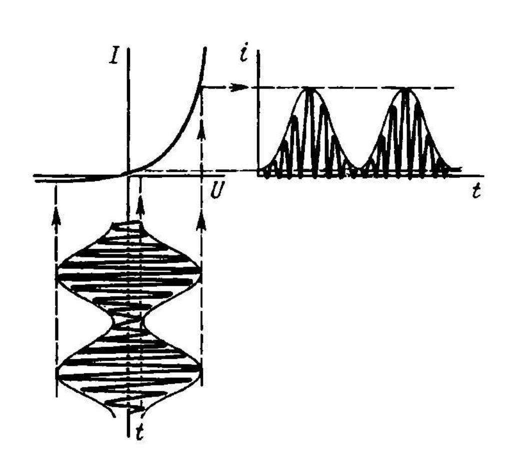


Рис. 3.52. Графики, поясняющие детектирование высокочастотных колебаний — выделение из модулированных по амплитуде высокочастотных колебаний сигнала более низкой частоты

Обобщенным параметром детекторного диода, учитывающим различные свойства диода и следующего за ним усилителя (видеоусилителя), является *коэффициент качества детекторного диода,* который характеризует чувствительность приемного устройства с детекторным диодом и определяется по формуле

$$M=\frac{\beta_I r_{\text{диф}}}{\sqrt{n_{\text{ш}} r_{\text{диф}}+r_{\text{ш}}}},$$

где $r_{\text{диф}}$ — дифференциальное сопротивление диода при определенном положительном смещении; $n_{\text{ш}}$ — шумовое отношение СВЧ-диода; $r_{\text{ш}}$ — эквивалентное шумовое сопротивление видеоусилителя, которое обычно принимают при расчетах равным 1 кОм.

Лучшие детекторные СВЧ-диоды имеют коэффициент качества более 100 $\text{Вт}^{-1/2}$. К таким диодам можно отнести, на-

пример, диоды Шотки с планарно-эпитаксиальной структурой на основе арсенида галлия АА204А...АА204В, предназначенные для детектирования в сантиметровом диапазоне длин волн.

Переключательные диоды

Переключательный полупроводниковый диод — это полупроводниковый диод, предназначенный для применения в устройствах управления уровнем сверхвысокочастотной мощности.

Принцип действия переключательного диода основан на большом различии полного сопротивления СВЧ-сигналу при прямом постоянном токе через диод и при обратном постоянном напряжении на диоде. Именно поэтому СВЧ-тракт (волноводная, коаксиальная или полосковая линия), следующий за переключательным устройством с диодом, может быть либо открыт, либо закрыт для СВЧ-сигнала. Например, в радиолокационных станциях с фазированными решетками, содержащими тысячи идентичных антенных элементов, переключательные диоды должны обеспечить подачу мощного СВЧ-импульса на каждый элемент в определенные моменты времени. При этом мощные импульсы передатчика не должны попадать в канал чувствительного приемника.

Отсюда ясны основные требования к переключательным СВЧ-диодам. Они должны с минимальными потерями пропускать СВЧ-мощность в состоянии пропускания и не пропускать — в состоянии запирания, обладать большой допустимой мощностью рассеяния, большим пробивным напряжением, малой собственной емкостью и достаточно большой скоростью переключения.

Обобщенным параметром переключательного диода является *критическая частота* $f_{кр}$, которая характеризует эффективность переключательного диода и определяется по формуле

$$f_{кр}=\frac{1}{2\pi C_{стр}\sqrt{r_{пр}r_{обр}}},$$

где $C_{стр}$ — емкость структуры; $r_{пр}$ — прямое сопротивление потерь (активная составляющая полного сопротивления диода) при определенном прямом токе смещения; $r_{обр}$ — обратное сопротивление потерь при определенном обратном напряжении смещения.

Для увеличения допустимой мощности рассеяния диода необходимо увеличивать площадь выпрямляющего электрического перехода, что влечет за собой увеличение барьерной емкости. Поэтому большинство переключательных СВЧ-диодов имеет *p-i-n*-структуру, толщина *p-n*-перехода которой существенно увеличена из-за наличия между *p*- и *n*-областями слоя высокоомного полупроводника с собственной электропроводностью (рис. 3.53).

Практически *p-i-n*-структуру для переключательных СВЧ-диодов формируют на исходном кристалле кремния с проводимостью, близкой к собственной, т. е. либо с небольшой концентра-

цией акцепторов (π-слой), либо с небольшой концентрацией доноров (ν-слой). Энергетическая диаграмма, распределение примесей, плотность объемного заряда и электрического поля в *p-i-n-* и *p-π-n*-структурах показаны на рис. 3.53. Методы формирования этих структур различны: вплавление и диффузия примесей, эпитаксиальное наращивание, ионное легирование.

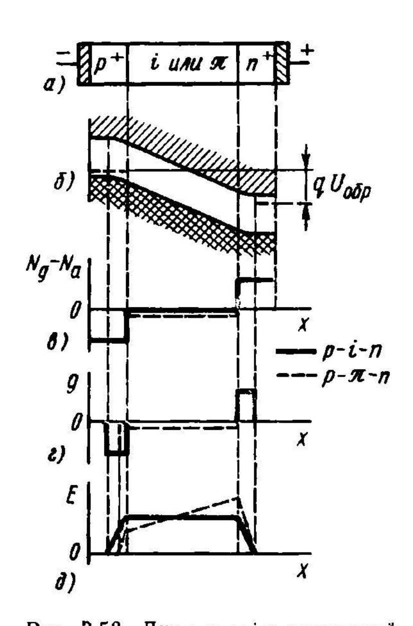


Рис. 3.53. Диод с *p-i-n*-структурой (*а*), энергетическая диаграмма (*б*), распределение примесей (*в*), плотности объемного заряда (*г*) и напряженности электрического поля (*д*)

Диоды с *p-i-n*-структурой отличаются меньшей барьерной емкостью, которая к тому же очень слабо зависит от напряжения (особенно при больших концентрациях примесей в *p-* и *n*-областях). Практическая независимость емкости структуры от напряжения оказывается важным свойством переключательных диодов, так как изменение емкости с напряжением может вызвать дополнительные частотные искажения полезного сигнала.

Пробивное напряжение диодов с *p-i-n*-структурой достигает нескольких сотен вольт, что существенно превышает пробивное напряжение диодов с обычным *p-n*-переходом и с таким же уровнем легирования прилегающих областей.

Для переключательных СВЧ-диодов некоторых марок (2А523А-4 и др.) максимально допустимая мощность, которую может рассеять диод в непрерывном режиме, равна 20 Вт. Такие диоды представляют собой бескорпусные приборы с жесткими выводами — кристаллодержателями — и защитным покрытием. Диаметр их 2 мм, длина 3,6 мм.

Переключательный СВЧ-диод может работать при последовательном и при параллельном включении с линией передачи. В параллельной схеме при прямом смещении диод имеет небольшое сопротивление, шунтирующее линию, и бо́льшая часть СВЧ-мощности отражается обратно. Таким образом, при параллельной схеме для переключения СВЧ-тракта используют разницу в отражении, а не в поглощении. В самом диоде при этом поглощается незначительная часть падающей на него СВЧ-мощности, что позволяет относительно маломощному прибору управлять десятками и сотнями киловатт импульсной СВЧ-мощности.

Недостатком переключательных СВЧ-диодов с *p-i-n*-структу-

рой является инерционность процесса рассасывания носителей заряда (электронов и дырок) из *i*-слоя при переключении диода с прямого направления на обратное, так как толщина *i*-слоя может составлять несколько десятков микрометров, а скорость движения носителей заряда ограничена.

Значительно большую скорость переключения можно получить при использовании диодов Шотки, изготовленных на основе арсенида галлия. Однако уровень переключаемой СВЧ-мощности при этом на несколько порядков ниже, чем при применении переключательных СВЧ-диодов с *p-i-n*-структурой.

§ 3.25. СТАБИЛИТРОНЫ

Полупроводниковый стабилитрон — это полупроводниковый диод, напряжение на котором в области электрического пробоя при обратном смещении слабо зависит от тока в заданном его диапазоне и который предназначен для стабилизации напряжения.

В стабилитронах должен быть либо лавинный, либо туннельный пробой, так как только при этих видах пробоя получаются ВАХ, необходимые для стабилизации напряжения (см. § 3.11, 3.12). В связи с тем что лавинный пробой характерен для диодов, изготовленных из полупроводника с большой шириной запрещенной зоны, исходным материалом для стабилитронов служит кремний.

Основным параметром стабилитронов является *напряжение стабилизации* $U_{ст}$ — значение напряжения на стабилитроне при прохождении заданного тока стабилизации (рис. 3.54). Пробивное напряжение диода, а значит, напряжение стабилизации стабилитрона, зависит от толщины *p-n*-перехода или от удельного сопротивления базы диода (см. рис. 3.21). Поэтому разные стабилитроны имеют различные напряжения стабилизации (от 3 до 400 В).

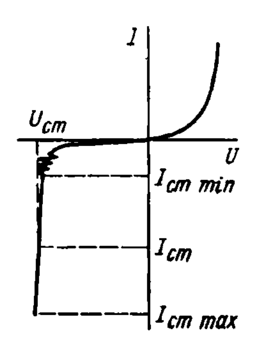


Рис. 3.54. ВАХ стабилитрона

Важным параметром стабилитрона является *температурный коэффициент напряжения стабилизации* $\alpha_{ст}$ — величина, определяемая отношением относительного изменения напряжения стабилизации к изменению температуры окружающей среды при постоянном токе стабилизации:

$$\alpha_{ст} = \frac{1}{U_{ст}} \frac{dU_{ст}}{dT} \bigg|_{I_{ст} = \text{const}}$$

Так как температурный коэффициент напряжения стабилизации зависит от температуры, то в справочной литературе приводят значения среднего температурного коэффициента напряжения стабилизации для рабочего диапазона температур:

$$\alpha_{ст} = \frac{1}{U_{ст.ср}} \frac{\Delta U_{ст}}{\Delta T}\Bigg|_{I_{ст}=\text{const}}.$$

Значения этого параметра у разных стабилитроиов различны. Обобщенная зависимость температуриого коэффициента напряжения стабилизации от напряжения стабилизации многих стабилитронов приведена иа рис. 3.55. Как видно из рисунка, $\alpha_{ст}$ может иметь положительиые значеиия для относительно высоковольтных и отрицательные для низковольтных стабилитроиов, что связано с различной температуриой зависимостью пробивного напряжения при лавиниом и туннельном пробое *p-n*-перехода (см. § 3.11, 3.12). Изменеиие знака $\alpha_{ст}$ соответствует иапряжению стабилизации $U_{ст} \approx 6$ В. Низковольтные стабилитроны изготовляют на основе сильнолегированного кремния. В связи с этим в низковольтиых стабилитронах с напряжеиием стабилизации менее 6 В происходит туииельиый пробой, а пробивное напряжение при туннельиом пробое уменьшается с увеличеиием температуры ($\alpha_{ст}<0$). Высоковольтные стабилитроиы должны иметь большую толщину *p-n*-перехода. Поэтому их делают иа основе слаболегированного кремния. Принцип их действия связан с лавинным пробоем, при котором пробивное напряжение растет с увеличением температуры ($\alpha_{ст}>0$). Среди различных типов стабилитронов большинство с лавинным пробоем.

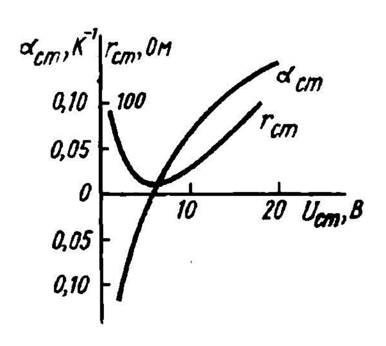


Рис. 3.55 Обобщенные зависимости температурного коэффициента напряжения стабилизации и дифференцнального сопротивления от напряжения стабилизации различиых стабилитроиов

Одним из способов уменьшения температурного коэффициента напряжения стабилизации, который используют для создаиия *термокомпенсированных* прецизионных стабилитронов, заключается в последовательном соедииении с обратио включенным *p-n*-переходом стабилитрона дополнительного *p-n*-перехода, включениого в прямом направлении. С повышением температуры напряжеиие на *p-n*-переходе, включенном в прямом иаправлении, уменьшается (см. § 3.2), что компенсирует увеличение напряжения на обратно включеном *p-n*-переходе при лавинном его пробое.

О качестве стабилитрона, т. е. о его способности стабилизировать напряжеиие при изменении проходящего тока, можно судить по значению *дифференциального сопротивления стабилитрона* $r_{ст}$, которое определяется отношением приращения напряжения стабилизации к вызвавшему его малому приращению тока. Так как определенным изменениям тока для лучшей стабилизации должны соответствовать минимальные изменеиия напряжения, то качество стабилитрона выше, если он имеет меньшее дифференциальное сопротивление.

Обобщенная зависимость дифференциального сопротивления от напряжения стабилизации многих стабилитронов показана на рис. 3.55. Для изготовления высоковольтных стабилитронов с лавинным пробоем в качестве исходного полупроводникового материала необходим высокоомный кремний. Чем больше требуется напряжение стабилизации, тем больше должно быть удельное сопротивление исходного кремния. При работе стабилитрона, т. е. при ударной ионизации в *p-n*-переходе, объемное сопротивление базы высоковольтного стабилитрона сказывается на значении дифференциального сопротивления. Поэтому с увеличением напряжения стабилизации дифференциальное сопро-

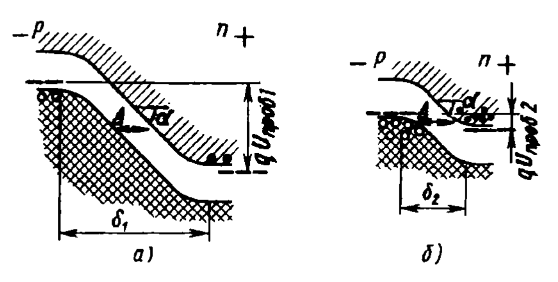


Рис. 3.56. Энергетические диаграммы, поясняющие увеличение дифференциального сопротивления с уменьшением напряжения стабилизации для стабилитронов с туннельным пробоем:

а — для стабилитрона с пробивным напряжением $U_{проб1}$; *б* — для стабилитрона с пробивным напряжением $U_{проб2} < U_{проб1}$

тивление высоковольтных стабилитронов увеличивается. Отсюда также ясна целесообразность формирования *p-n*-переходов высоковольтных стабилитронов в тонком высокоомном эпитаксиальном слое кремния, выращенном на низкоомной подложке.

Для низковольтных стабилитронов с туннельным пробоем при увеличении концентрации примесей уменьшается толщина *p-n*-перехода, что и приводит к уменьшению пробивного напряжения и напряжения стабилизации. Рассмотрим теперь причины увеличения дифференциального сопротивления низковольтных стабилитронов с уменьшением их напряжения стабилизации (левая часть графика $r_{ст} = f(U_{ст})$ на рис. 3.55).

Для сравнения на рис. 3.56 приведены энергетические диаграммы двух низковольтных стабилитронов с разными напряжениями стабилизации. Для туннелирования носителей заряда сквозь *p-n*-переход, во-первых, должна быть мала толщина Δ потенциального барьера, сквозь который должны туннелировать электроны. Эта толщина потенциального барьера получится при определенной напряженности электрического поля или угле наклона энергетических зон, так как $E \sim \text{tg}\,\alpha$. Во-вторых, для туннелирования необходимо наличие электронов с одной стороны *p-n*-перехода (в данном случае в валентной зоне *p*-области)

и свободных энергетических уровней, соответствующих тем же значеииям эиергии, с другой стороны *p-n*-перехода (в данном случае в зоие проводимости *n*-области). Оба эти условия возиикли в двух сравниваемых стабилитронах при разиых пробивиых иапряжениях. Но иепосредствеиио под потолком валентиой зоиы *p*-области электроиов значительио меиьше, чем иа более глубоких уровнях валентиой зоиы. Особеиио это справедливо для второго из сравииваемых стабилитроиов, концентрация примесей в котором значительио больше. Поэтому с дальиейшим приращением обратного напряжения приращение числа электронов, способиых протуинелировать, в первом стабилитроие оказывается зиачительно большим, чем во втором. Следовательио, дифференциальиое сопротивлеиие первого стабилитроиа должио быть меньше, чем второго.

Диапазон токов, в котором стабилитрон может выполнять фуикции стабилизации иапряжения, устаиавливают из следующих соображений. *Минимально допустимый ток стабилизации* $I_{\text{ст min}}$ определяется тем, что при малых токах, во-первых, днфферепциальное сопротивление оказывается еще большим и, во-вторых, в стабилитронах с лавинным пробоем из-за неустойчивости процесса удариой иоиизации в иачальной стадии возникают шумы. С увеличением тока через стабилитроны процесс удариой ионизации устаиавливается и шумы исчезают. *Максимально допустимый ток стабилизации* $I_{\text{ст max}}$ определяется допустимой для даииого прибора мощностью рассеяиия и иеобходимостью обеспечеиия заданной иадежности работы прибора, т. е. зависнт от площади *p-n*-перехода и от коиструкции прибора.

Конструктивио стабилитроны оформляют аналогично выпрямительным и другим диодам, т. е. в металлостеклянных, стеклянных и пластмассовых корпусах, а также в бескорпусном виде с защитиым покрытием. В зависимости от площади *p-n*-перехода и конструктивиого оформлеиия стабилитроны могут иметь максимальио допустимые мощиости от десятых долей до 50 Вт.

Электроино-дырочные переходы при изготовлеиии стабилитроиов формируют методами вплавлеиия и диффузии примесей. При вплавлении или при диффузии одной и той же примеси с двух сторон кристалла кремиия можно сформировать одиовременно два *p-n*-перехода, которые при подаче напряжеиия иа крайиие области структуры окажутся включеиными встречно. Так изготовляют стабилитроны с симметричиой ВАХ — *двуханодные стабилитроны,* предиазначеиные для применения в схемах стабилизации иапряжеиия разиой поляриости и для защиты различных элементов электрических цепей от перенапряжений обеих полярностей.

§ 3.26. СТАБИСТОРЫ

Полупроводниковый стабистор — это полупроводниковый диод, напряжение на котором в области прямого смещения слабо зависит от тока в заданном его диапазоне и который предназначен для стабилизации напряжения.

Отличительной особенностью стабисторов по сравнению со стабилитронами является меньшее напряжение стабилизации, определяемое прямым падением напряжения на диоде, и составляет примерно 0,7 В. Последовательное соединение двух или трех стабисторов дает возможность получить удвоенное или утроенное значение напряжения стабилизации. Некоторые типы стабисторов представляют собой единый прибор с последовательным соединением отдельных элементов.

Стабисторы имеют отрицательный температурный коэффициент напряжения стабилизации, т. е. напряжение на стабисторе при неизменном токе уменьшается с увеличением температуры. Связано это, во-первых, с уменьшением высоты потенциального барьера на *p-n*-переходе при увеличении температуры (см. § 2.1) и, во-вторых, с перераспределением носителей заряда по энергиям, что с увеличением температуры приводит к переходу через потенциальный барьер большего числа носителей. В связи с отрицательным температурным коэффициентом напряжения стабилизации и нелинейностью ВАХ, которая обеспечивает стабилизацию напряжения, стабисторы используют для температурной компенсации стабилитронов с положительным температурным коэффициентом напряжения стабилизации. Для этого последовательно со стабилитроном необходимо соединить один или несколько стабисторов.

Основная часть стабисторов — это кремниевые диоды, отличающиеся от обычных выпрямительных диодов тем, что *p-n*-переходы для стабисторов формируют в низкоомном кремнии. Это необходимо для получения меньшего объемного сопротивления базы и соответственно меньшего дифференциального сопротивления стабистора. Сопротивление базы может влиять на значение дифференциального сопротивления стабистора, так как его *p-n*-переход при работе смещен в прямом направлении и имеет малое сопротивление.

Из-за малого удельного сопротивления исходного кремния толщина *p-n*-перехода и пробивное напряжение стабисторов оказываются очень малыми, но стабисторы предназначены для работы при прямом включении. Обратное напряжение на них может оказаться только во время переходных процессов в той или иной схеме. Максимально допустимое обратное напряжение при переходных процессах обычно не превышает для стабисторов нескольких вольт.

Кроме кремниевых стабисторов промышленность выпускает и селеновые поликристаллические стабисторы, которые отличаются простотой изготовления, а значит, меньшей стоимостью. Однако

селеновые стабисторы имеют меньший гарантированный срок службы (1000 ч) и узкий диапазон рабочих температур (−25...+60°C).

§ 3.27. ШУМОВЫЕ ДИОДЫ

Полупроводниковый шумовой диод — это полупроводниковый прибор, являющийся источником шума с заданной спектральной плотностью в определенном диапазоне частот.

В начальной стадии лавинного пробоя, как отмечено в § 3.25, процесс ударной ионизации оказывается неустойчивым: ударная ионизация возникает, срывается, возникает вновь в тех местах *p-n*-перехода, где оказывается в данный момент достаточная напряженность электрического поля. Результатом случайной неравномерности генерации новых носителей заряда при ударной ионизации являются шумы, которые характерны для определенного диапазона токов (см. рис. 3.54). При работе таких, например, приборов, как стабилитроны, шумы — явление вредное. Именно поэтому диапазон токов, соответствующий шумам, исключают из диапазона рабочих токов стабилитронов. Однако для различных измерений в радиотехнике нужны генераторы шумовых напряжений.

Таким образом, в качестве генератора шумовых напряжений можно использовать диод в диапазоне обратных токов от минимального $I_{\text{проб min}}$ до максимального $I_{\text{проб max}}$ пробивного тока, где наблюдается наибольшая интенсивность электрических флуктуаций. Так, для шумовых диодов КГ401А...КГ401В этот диапазон соответствует значениям токов 10 мкА...1 мА.

Основными параметрами шумовых диодов являются *спектральная плотность шума* $S_{\text{ш}}$ — эффективное значение напряжения шума, отнесенное к 1 Гц, при заданном токе пробоя, а также *граничная частота равномерности спектра* $f_{\text{гр}}$ — наибольшая частота спектра, при которой удовлетворяется в сторону отрицательного отклонения заданное требование по неравномерности спектральной плотности шума (при заданном токе пробоя).

Одним из справочных параметров шумовых диодов является *средний температурный коэффициент спектральной плотности шума* ТК $S_{\text{ш}}$ — отношение относительного изменения спектральной плотности шума в заданном диапазоне рабочих температур к абсолютному изменению температуры окружающей среды при постоянном токе:

$$\text{ТК}\, S_{\text{ш}} = \frac{1}{S_{\text{ш ср}}} \frac{\Delta S}{\Delta T} \bigg|_{I_{\text{проб}} = \text{const}} .$$

Обратный ток, предшествующий лавинному пробою, и пробивное напряжение при лавинном пробое увеличиваются с ростом температуры. В результате участок ВАХ, соответствующий наибольшей интенсивности шумов, смещается с изменением температуры в область больших токов и напряжений. Поэтому знак

и значение температурного коэффициента спектральной плотности шума могут быть различными при разных постоянных токах, при которых производят измерение ТК $S_ш$ шумового диода.

§3.28. ЛАВИННО-ПРОЛЕТНЫЕ ДИОДЫ

Лавинно-пролетный диод — это полупроводниковый диод, работающий в режиме лавинного размножения носителей заряда при обратном смещении электрического перехода и предназначенный для генерации сверхвысокочастотных колебаний.

Генерация электромагнитных СВЧ-колебаний может возникать в диодах с различной структурой. В качестве примера рассмотрим процессы, происходящие в структуре p^+-n-n^+ при обратном напряжении, имеющем постоянную и переменную составляющие. Когда суммарное напряжение превышает пробивное, начинается ударная ионизация — лавинный пробой. Пары электрон — дырка, генерируемые в узкой части p-n-перехода

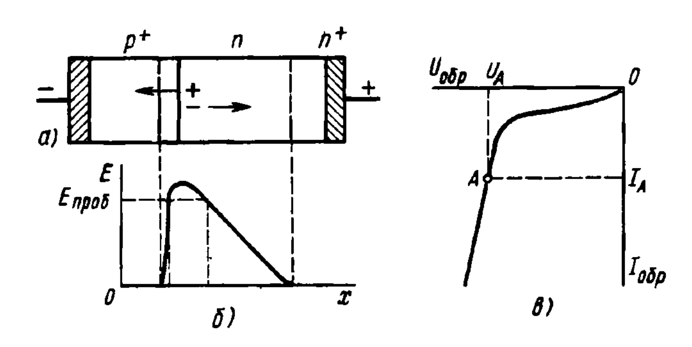


Рис. 3.57. Структура лавинно-пролетного диода (*а*), распределение напряженности электрического поля по структуре (*б*) и положение рабочей точки (постоянного смещения) на ВАХ (*в*)

вблизи металлургической границы, где напряженность электрического поля достаточна для ударной ионизации, разделяются полем (рис. 3.57). Ток, вызванный движением новых носителей заряда, проходит до тех пор, пока эти носители на выйдут из p-n-перехода. За время пролета носителей заряда через переход (в нашем примере — электронов) напряжение на диоде может успеть уменьшиться, если частота переменной составляющей будет большой. Таким образом, из-за конечного времени пролета носителей появляется фазовый сдвиг между проходящим через диод током и приложенным к этому диоду переменным напряжением высокой частоты.

Фазовый сдвиг между током и напряжением определяется не только временем пролета, но и инерционностью процесса

развития лавины при ударной ионизации. Действительно, момент времени приобретения носителем заряда достаточной для ионизации энергии, вероятнее всего, не совпадает с моментом столкновения этого носителя с одним из атомов полупроводника, т. е. с моментом ионизации. Кроме того, для приобретения добавочной энергии также необходимо некоторое время.

Допустим, что время пролета вместе с временем, определяемым инерционностью ударной ионизации, равно половине периода колебаний некоторой частоты переменного напряжения (рис. 3.58, *а*). В этом случае переменный ток через диод будет отставать на полпериода от вызвавшего его переменного напряжения. Нарастание напряжения все время будет сопровож-

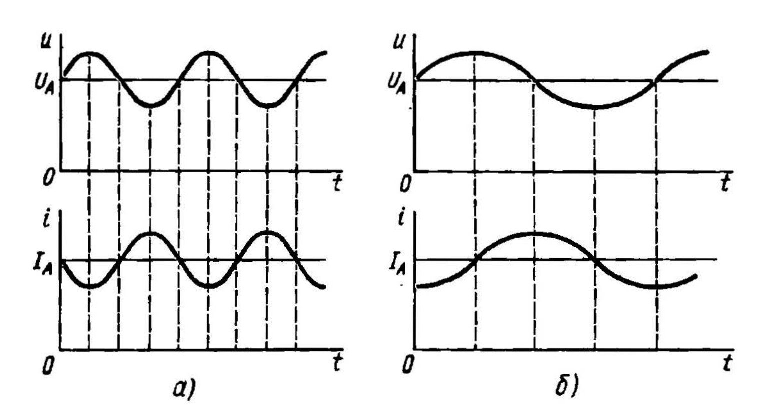


Рис. 3.58. Зависимости напряжений и токов, иллюстрирующие появление отрицательного дифференциального сопротивления лавинно-пролетного диода:

а — при сдвиге фаз 180°, возникающем на больших частотах переменной составляющей напряжения; *б* — при сдвиге фаз 90°, соответствующем меньшим частотам переменной составляющей напряжения и отсутствию отрицательного дифференциального сопротивления в среднем за период

даться уменьшением тока, а уменьшение напряжения, наоборот, ростом тока. Это свидетельствует о том, что для данной частоты переменного напряжения в течение всего периода колебаний выполняется условие отрицательного дифференциального сопротивления.

При понижении частоты переменного напряжения (при увеличении периода колебаний) ток будет отставать от напряжения на угол, меньший 180°, так как время пролета и инерционность ударной ионизации остаются теми же. Когда с понижением частоты переменного напряжения фазовый сдвиг между током и напряжением будет составлять четверть периода, условия отрицательного дифференциального сопротивления будут выполняться на протяжении только половины периода, чередуясь через каждые четверть периода с условиями положительного дифференциального сопротивления (рис. 3.58, *б*). В этом предельном

случае в среднем за период лавинно-пролетный диод не будет обладать отрицательным дифференциальным сопротивлением.

Аналогично с повышением частоты переменного напряжения отрицательное дифференциальное сопротивление пропадает при фазовом сдвиге между током и напряжением, достигающем 270°. Таким образом, лавинно-пролетные диоды обладают отрицательным дифференциальным сопротивлением только для СВЧ-колебаний.

Всякий прибор с отрицательным дифференциальным сопротивлением может быть использован для генерации и усиления электромагнитных колебаний. Лавинно-пролетные диоды используют для генерации СВЧ-колебаний большой мощности. При этом не обязательно подавать на лавинно-пролетный диод переменное напряжение необходимой частоты, что мы мысленно делали при рассмотрении принципа действия прибора. Лавинно-пролетный диод вместе с резонансной камерой, в которую его обычно помещают, способен выделить из импульсов, возникающих во время подачи постоянного смещения, и усилить колебания определенной частоты.

Рис. 3.59. Распределение напряженности электрического поля в слаболегированной n-области p-n-перехода в различные моменты времени при работе лавинно-пролетного диода в режиме с захваченной плазмой

Лавинно-пролетные диоды впервые были созданы в СССР на основе открытого А. С. Тагером с сотрудниками эффекта генерации и усиления СВЧ-колебаний при лавинном пробое.

Кроме рассмотренного *лавинно-пролетного режима* работы, который в английской литературе называют IMPATT-режимом (сокращение от impact ionization avalanche transit time), лавинно-пролетные диоды могут работать и в *режиме с захваченной плазмой* или TRAPATT-режиме (trapped plasma avalanche triggered transit).

Принцип действия при этом режиме работы связан с тем, что скорость перераспределения электрического поля в структуре диода может значительно превышать скорость дрейфа носителей заряда. На рис. 3.59 показано распределение напряженности электрического поля в слаболегированной n-области p^+-n-n^+-структуры лавинно-пролетного диода в различные моменты времени после включения диода на обратное напряжение, превышающее пробивное напряжение. В первый момент (t_1) напряженность электрического поля максимальна около металлургической границы. Именно здесь из-за ударной ионизации начинается образование электронно-дырочной плазмы. Это приводит к перераспределению электрического поля в n-области.

В следующий момент времени (t_2) ударная ионизация будет происходить в соседнем слое *n*-области. Скорость дрейфа носителей заряда ограничеиа даже в сильных электрнческих полях скоростью насыщеиия (см. рис. 1.13). Скорость дрейфа электронов плазмы может оказаться значительно меньше скорости насыщения, если напряженность электрического поля в слое с плазмой успеет уменьшиться. Перераспределение напряженности электрического поля может произойти быстрее, если источннк питания диода обеспечит большую плотность тока через диод (с учетом плотности тока смещеиия), что подтверждает формула (1.32), и если концентрация примесей в слаболегированной области будет достаточна мала.

В результате фронт волны ионизации быстро пересекает всю *n*-область, которая заполняется высокопроводящей электронно-дырочной плазмой. Напряженность электрического поля в это время (t_3 на рис. 3.59) и напряжение на диоде становятся малыми, что приводит к относительно медленному рассасыванию носнтелей плазмы из *p-n*-перехода. Задержка экстракции носителей из *p-n*-перехода обусловила название «режим с захваченной плазмой».

Так как скорость направленного движения носителей заряда в лавинно-пролетных диодах в режиме с захваченной плазмой значительно ниже скорости насыщения, то частота генерируемых колебаний обычно не превышает 10 ТГц, в то время как при лавинно-пролетном режиме эта частота может составлять несколько сотен гигагерц. Другие отличия в свойствах и параметрах при различных режимах работы вызваны тем, что при лавинно-пролетном режиме уменьшение скорости дрейфа ниже скорости насыщения нежелательно, а при режиме с захваченной плазмой — наоборот. Поэтому бóльшая амплитуда колебаний может быть получена в режиме с захваченной плазмой — до нескольких сотен киловатт при импульсиой работе (при непрерывной работе — до нескольких ватт). А так как напряжение на диоде оказывается малым при больших токах и, наоборот, большим при малых токах, то КПД достигает 40% для диодов из арсенида галлия и кремния.

Лавинно-пролетные диоды имеют высокий уровеиь шума, присущий ударной ионизации, так как небольшие беспорядочные отклонения тока (дробовóй шум) через электрический переход усиливаются при ударной ионизации в число раз, равное коэффициенту лавинного размножения *M*. Поэтому лавинно-пролетные диоды применяют только для генерации СВЧ-колебаннй и не используют для усилення слабых сигналов. Поэтому же лавинный пробой используют для создания шумовых диодов.

§ 3.29. ТУННЕЛЬНЫЕ ДИОДЫ

Структура и принцип действия

Туннельный диод — это полупроводниковый диод на основе вырожденного полупроводника, в котором туннельный эффект приводит к появлению на вольт-амперной характеристике при прямом напряжении участка отрицательной дифференциальной проводимости.

В отличие от всех остальных полупроводниковых диодов для изготовления туннельных диодов используют полупроводниковый материал с очень высокой концентрацией примесей ($10^{18}...10^{20}$ см$^{-3}$). Следствием высокой концентрации примесей в прилегающих к *p-n*-переходу областях является, во-первых,

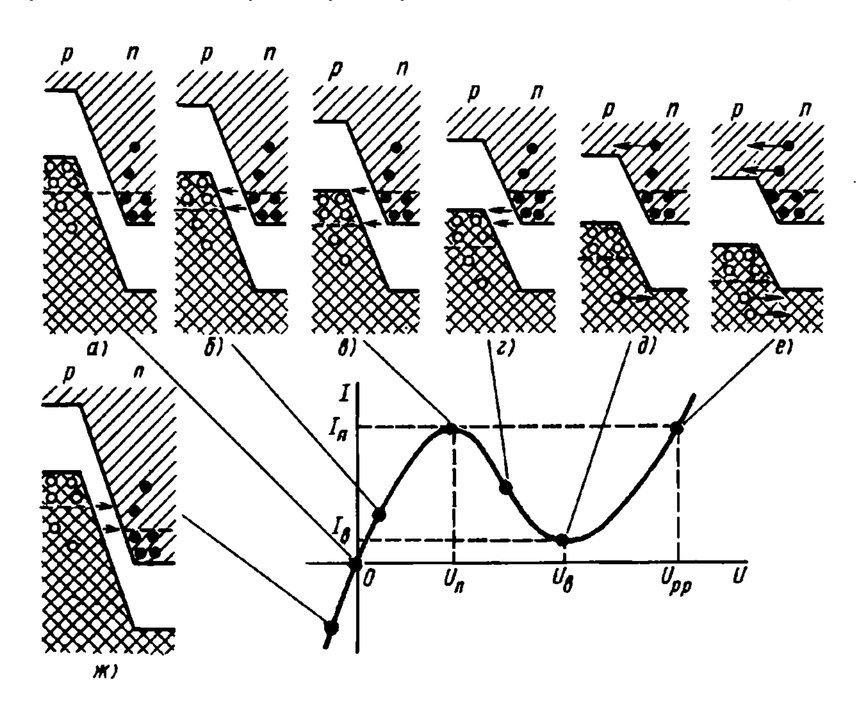


Рис. 3.60. ВАХ и энергетические диаграммы туннельного диода при: *а* — отсутствии напряжения; *б* —небольшом прямом напряжении; *в* — пиковом напряжении; *г* — напряжении, соответствующем отрицательному дифференциальному сопротивлению; *д* — напряжении впадины; *е* — напряжении раствора, вызывающем значительный инжекционный ток; *ж* — обратном напряжении

малая толщина перехода (около 10^{-2} мкм), т. е. на два порядка меньше, чем в других полупроводниковых диодах. Сквозь такие тонкие потенциальные барьеры возможно туннелирование носителей заряда (см. § 1.10).

Другим следствием большой концентрации примесей является расщепление примесных энергетических уровней с образова-

нием примесных энергетических зон, которые примыкают к зоне проводимости в *n*-области и к валентной зоне в *p*-области. Уровни Ферми при этом оказываются расположенными в разрешенных зонах (рис. 3.60).

В диоде без внешнего напряжения существует туннелирование электронов из *n*-области в *p*-область и обратно. Встречные потоки электронов равны, поэтому суммарный ток через диод равен нулю (рис. 3.60, *а*).

При небольшом прямом напряжении на туннельном диоде происходит уменьшение высоты потенциального барьера *p-n*-перехода или смещение энергетической диаграммы *n*-области относительно энергетической диаграммы *p*-области. Свободные энергетические уровни *p*-области (занятые дырками), расположенные непосредственно над уровнем Ферми, оказываются на одной высоте по энергетической диаграмме или при одних и тех же значениях с энергетическими уровнями *n*-области, занятыми электронами (рис. 3.60, *б*). Поэтому будет происходить преимущественное туннелирование электронов из *n*-области в *p*-область.

При прямом напряжении на диоде, когда свободные энергетические уровни валентной и примесной зон *p*-области окажутся на одной высоте с занятыми электронами энергетическими уровнями зоны проводимости и примесной зоны *n*-области, туннельный ток через диод будет максимальным (рис. 3.60, *в*).

При дальнейшем увеличении прямого напряжения на диоде туннельный ток через диод будет уменьшаться, так как из-за смещения энергетических диаграмм будет уменьшаться количество электронов, способных туннелировать из *n*-области в *p*-область (рис. 3.60, *г*).

Туннельный ток через диод окажется равным нулю при некотором еще большем прямом напряжении, когда из-за относительного смещения энергетических диаграмм *n*- и *p*-областей для свободных электронов *n*-области не будет свободных энергетических уровней в *p*-области (рис. 3.60, *д*). Однако при этом через диод будет проходит прямой ток, обусловленный переходом носителей заряда через понизившийся потенциальный барьер *p-n*-перехода, т. е. ток, связанный с инжекцией.

С дальнейшим увеличением прямого напряжения в связи с уменьшением высоты потенциального барьера прямой ток через туннельный диод будет возрастать, как и в обычных выпрямительных диодах (рис. 3.60, *е*).

При обратном напряжении на туннельном диоде опять возникают условия для туннелирования электронов (рис. 3.60, *ж*). Только теперь электроны туннелируют из валентной зоны *p*-области в зону проводимости *n*-области. Возникающий при этом обратный ток будет расти с увеличением обратного напряжения по абсолютному значению. Туннельный диод обладает относительно высокой проводимостью при обратном напряжении. Можно считать, что у туннельного диода при ничтожно малых обратных напряжениях происходит туннельный пробой.

Таким образом, туннельный диод обладает отрнцательным дифференциальным сопротивлением в некотором диапазоне прямых напряжений. Это и является самым интересным свойством туннельного диода, так как всякий прибор с отрицательным дифференциальным сопротивлением может быть использован для генерации и усиления электромагнитных колебаний, а также в переключающих схемах.

Параметры

Промышленность выпускает туннельные диоды из арсенида галлия и из германия. Их вольт-амперные характеристики показаны на рис. 3.61. Из рисунка видно, что чем больше шириια запрещенной зоны исходного полупроводника, тем при больших напряжениях наблюдается отрицательное дифференциальиое сопротивление.

Туннельные диоды характеризуются специфнческими параметрами (см. рис. 3.60):

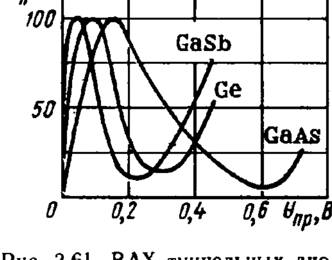


Рис. 3.61. ВАХ туннельных диодов из различных полупроводниковых материалов

1. *Пиковый ток* $I_П$ — прямой ток в точке максимума ВАХ, при котором значение di/du равно нулю. Этот ток различен для туннельных диодов разного назначения. Значение его может составлять от десятых долей миллиампера до сотен миллиампер.

2. *Ток впадины* $I_В$ — прямой ток в точке миннмума ВАХ, при котором значение di/du равно нулю.

3. *Отношение токов туннельного диода* $I_П/I_В$ — отношение пикового тока к току впадины. Для туннельных диодов из арсенида галлия $I_П/I_В \geqslant 10$, для германиевых туннельных диодов $I_П/I_В = 3 ... 6$.

4. *Напряжение пика* $U_П$ — прямое напряжение, соответствующее пиковому току. Для туннельных диодов из арсенида галлия $U_П = 100...150$ мВ, для германиевых $U_П = 40...60$ мВ.

5. *Напряжение впадины* $U_В$ — прямое напряжение, соответствующее току впадины. У туннельных диодов из арсенида галлия $U_В = 400...500$ мВ, у германиевых $U_В = 250...350$ мВ.

6. *Напряжение раствора* $U_{РР}$ — прямое напряжение, большее напряження впадины, при котором ток равен пиковому.

7. *Удельная емкость туннельного диода* $C_д/I_п$ — отношение емкости туннельного диода к пиковому току.

8. *Предельная резистивная частота* f_R — расчетная частота, на которой активная составляющая полного сопротивления последовательной цепи, состоящей из *p-n*-перехода и сопротивления потерь, обращается в нуль.

9. *Резонансная частота туннельного диода* f_0 — расчетная частота, на которой общее реактивное сопротивление *p-n*-перехода и индуктивности корпуса туннельного диода обращается в нуль.

Температурные зависимости параметров

Изменение температуры туннельного диода может по-разному влиять на туннельную составляющую тока и на составляющую, связанную с инжекцией.

На температурную зависимость туннельной составляющей тока могут влиять следующие физические факторы.

1. С повышением температуры уменьшается ширина запрещенной зоны арсенида галлия и германия — основных исходных полупроводниковых материалов для туннельных диодов. Уменьшение ширины запрещенной зоны приводит к уменьшению толщины потенциального барьера, сквозь который туннелируют электроны, при этом вероятность туннелирования растет. Туннельная составляющая тока и, в частности, пиковый ток увеличиваются.

2. При увеличении температуры изменяется распределение электронов по энергетическим уровням — количество электронов под уровнем Ферми в зоне проводимости *n*-области уменьшается, так как часть свободных электронов переходит на более высокие энергетические уровни, а уровень Ферми смещается вниз. Поэтому уменьшается число электронов, которые могут туннелировать из *n*-области в *p*-область. Туннельная составляющая прямого тока уменьшается.

Так как эти факторы действуют, так сказать, в разные стороны, то суммарное их влияние, во-первых, должно быть малым, а во-вторых, оно может привести как к увеличению, так и к уменьшению пикового тока туннельного диода с увеличением температуры.

Инжекционная составляющая тока туннельного диода растет с увеличением температуры по двум причинам, имеющим место и в выпрямительных диодах (см. § 3.2): уменьшение высоты потенциального барьера и перераспределение носителей заряда по энергетическим уровням. Поэтому у туннельного диода ток впадины растет с увеличением температуры.

Частотные свойства туннельных диодов

Механизм действия туннельных диодов связан с туннелированием электронов сквозь потенциальный барьер. Время, необходимое для завершения этого процесса, составляет $10^{-13}...10^{-14}$ с. Эффекта накопления неосновных носителей в базе туннельных диодов практически нет, так как они используются при малых

напряжениях, соответствующих падающему участку ВАХ (с отрицательиым дифференциальным сопротивлением). Поэтому туннельные диоды способны работать на частотах до сотен гигагерц, что соответствует миллиметровому диапазону радиоволн. Верхний предел частотиого диапазона работы туинельных диодов ограиичен лишь паразитными реактивностями — собственной емкостью, основу которой составляет барьерная емкость *p-n*-перехода, и индуктивностью выводов и корпуса.

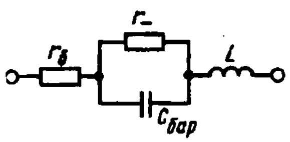


Рис. 3.62. Эквивалентная схема туннельного диода

Для анализа и расчета параметров, характеризующих частотные свойства туннельных диодов, воспользуемся эквивалентной схемой туннельного диода для малого переменного сигнала при наличии постоянного напряжения, которое смещает рабочую точку на падающий участок ВАХ. Эквивалентная схема туннельного диода отличается от эквивалентной схемы обычного диода только тем, что здесь вместо активного сопротивления перехода введено отрнцательное дифференциальное сопротивление r_- и учитывается индуктивность выводов L (рис. 3.62).

Полное сопротивление туннельного днода при малом синусоидальном напряжении

$$\dot{z} = r_б + j\omega L + \frac{1}{g + j\omega C_{бар}} = r_б + \frac{g}{g^2 + \omega^2 C_{бар}} + $$
$$+ j\left(\omega L - \frac{\omega C_{бар}}{g^2 + \omega^2 C_{бар}^2}\right). \qquad (3.104а)$$

где $g = 1/r_-$ — отрицательная дифференцнальная проводимость туннельного днода.

Отрицательное днфференцнальное сопротивление у туниельного днода будет существовать, если вещественная часть полиого сопротнвлення будет меньше нуля, т. е. при

$$r_б < \left|\frac{g}{g^2 + \omega^2 C_{бар}^2}\right|. \qquad (3.105)$$

При большой частоте второе слагаемое в вещественой части (3.104а) окажется меньше (по абсолютному значению) первого слагаемого. Это значит, что туннельный диод при больших частотах не будет обладать отрицательным дифференциальным сопротивленнем. Определим *предельную резистивную частоту,* на которой активная составляющая полного сопротнвления туннельного диода обращается в нуль, т. е. неравенство (3.105) обращается в равенство

$$r_б = \frac{g}{g^2 + \omega_R^2 C_{бар}^2}.$$

Отсюда

$$\omega_R = \frac{1}{|r_-| C_{бар}} \sqrt{\frac{|r_-|}{r_б} - 1}\,. \qquad (3.106)$$

Из соотиошеиия (3.106) видно, что предельиая резистивная частота туннельного днода зависит от соотиошения сопротивления базы диода и абсолютиого значения отрицательного дифференцнального сопротивления. Для нахождеиия экстремальных значеннй предельной резистивной частоты иадо проднфференцировать (3.106) по $|r_-|$ н приравиять полученное выражение иулю. Тогда получнм условие

$$|r_-| = 2r_б\,, \qquad (3.107)$$

при выполиении которого предельная резистивная частота будет максимальной. Связь ее с параметрами эквивалентиой схемы можио получить, подставнв (3.107) в (3.106):

$$f_R = \frac{1}{2\pi |r_-| C_{бар}}\,. \qquad (3.108)$$

Усредненное отрицательиое дифференциальиое сопротнвление можно представить в виде

$$|r_-| = \frac{\Delta U}{\Delta I} = \frac{U_В - U_П}{I_П - I_В} \approx \frac{\mathrm{const}}{I_П}\,.$$

Если полученное соотношеине подставить в (3.108), то получнм

$$f_R \sim I_П / C_{бар}\,.$$

Поэтому отношеиие емкостн туннельного диода к пиковому току является одним из основиых параметров туннельного днода, который характеризует его частотные свойства.

В туннельном диоде, имеющем паразитиую индуктивиость и емкость, при некоторой частоте могут возникнуть условия для резонанса, что нарушит нормальную работу схемы с туииельным днодом. Поэтому резонансиая частота туннельного диода f_0 является еще одним параметром, характеризующим частотные свойства туннельного диода.

Резонансную частоту туииельного диода можно определить из соотношеиия (3.104а) при условии равеиства нулю миимой части полного сопротивлеиия диода:

$$\omega_0 L = \frac{\omega_0 C_{бар}}{g^2 + \omega_0^2 C_{бар}^2}\,,$$

тогда

$$\omega_0 = \frac{1}{|r_-| C_{бар}} \sqrt{\frac{|r_-|^2 C_{бар}}{L} - 1}\,. \qquad (3.109)$$

Туннельиые диоды должиы быть сконструироваиы так, чтобы резонаисная частота была выше предельной резистивной частоты

($\omega_0 > \omega_R$), тогда возможные нежелательные резонансные явления возникают только на частотах, на которых туннельный диод уже не будет обладать отрицательным дифференциальным сопротивлением. Неравенство $\omega_0 > \omega_R$ с учетом соотношений (3.109) и (3.106) преобразуем следующим образом: $L < r_б \mid r_- \mid C_{бар}$.

Следовательно, индуктивность эквивалентной схемы туннельного диода, которая определяется в основном индуктивностью внутренних и внешних выводов диода, должна быть по возможности малой. Поэтому внутренние и внешние выводы у туннельных диодов делают не из тонкой проволоки, а из плющенки или в виде мембран, пластин и т. п.

§ 3.30. ОБРАЩЕННЫЕ ДИОДЫ

Обращенным называют диод на основе полупроводника с критической концентрацией примесей, в котором проводимость при обратном напряжении вследствие туннельного эффекта значительно больше, чем при прямом напряжении.

При концентрациях примесей в *p*- и *n*-областях диода, меньших, чем в туннельных диодах, но больших, чем в обычных выпрямительных диодах, можно получить диод, энергетическая диаграмма которого показана на рис. 3.63, *а*. Уровень Ферми при такой сред-

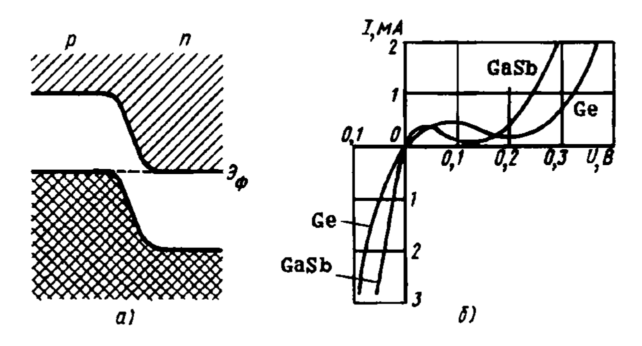


Рис. 3.63. Энергетическая диаграмма (*а*) и ВАХ (*б*) обращенных диодов

ней концентрации примесей может быть расположен на потолке валентной зоны *p*-области и на дне зоны проводимости *n*-области диода, т. е. потолок валентной зоны *p*-области и дно зоны проводимости *n*-области при нулевом смещении на диоде находятся на одной высоте по энергетической диаграмме.

Обратная ветвь ВАХ обращенного диода аналогична обратной ветви ВАХ туннельного диода, так как при обратных напряжениях происходит туннелирование электронов из валентной зоны *p*-области в зону проводимости *n*-области. Поэтому обратные токи в

обращенных диодах оказываются большими при ничтожно малых обратных напряжениях (десятки милливольт).

Прямая ветвь ВАХ обращенного диода аналогична прямой ветви ВАХ обычного выпрямительного диода, так как при прямых напряжениях на обращенном диоде прямой ток может быть образован только в результате инжекции носителей заряда через потенциальный барьер *p-n*-перехода. Но заметная инжекция может наблюдаться только при прямых напряжениях в несколько десятых долей вольта. При меньших напряжениях прямые токи в обращенных диодах оказываются меньше обратных (рис. 3.63, *б*).

Таким образом, обращенные диоды обладают выпрямляющим эффектом, но пропускное (проводящее) направление у них соответствует обратному включению, а запирающее (непроводящее) — прямому включению.

Из принципа действия обращенных диодов ясно, что они, во-первых, способны работать на очень малых сигналах. Во-вторых, должны обладать хорошими частотными свойствами, так как туннелирование — процесс малоинерционный, а эффекта накопления неосновных носителей при малых прямых напряжениях практически нет. Поэтому обращенные диоды можно использовать на СВЧ. В-третьих, из-за относительно большой концентрации примесей в прилегающих к *p-n*-переходу областях обращенные диоды оказываются мало чувствительными к воздействиям проникающей радиации.

§ 3.31. ВАРИКАПЫ

Варикап — это полупроводниковый диод, действие которого основано на использовании зависимости емкости от обратного напряжения и который предназначен для применения в качестве элемента с электрически управляемой емкостью.

Как было показано в § 2.8 и 3.6, диод обладает барьерной и диффузионной емкостями. В качестве варикапов используют только диоды при обратном постоянном смещении, когда проявляется только барьерная емкость. Диффузионная емкость проявляется при прямом смещении диода, когда проводимость его велика и велики потери мощности из-за относительно больших активных токов через диод.

Зависимость емкости от напряжения смещения различна для варикапов, изготовленных методом диффузии или методом вплавления примесей (см. § 2.8). В сплавных варикапах с резким *p-n*-переходом зависимость барьерной емкости от напряжения смещения получается более резкая. Связано это с тем, что глубина проникновения электрического поля в материал зависит от удельного сопротивления этого материала. В сплавном варикапе слои базы, прилегающие к переходу, легированы равномерно, а в диффузионном — при удалении от перехода концентрация нескомпенсированных примесей увеличивается, т. е. уменьшается удельное сопротивление.

Поэтому для получения еще более резкой зависимости емкости

варикапа от напряжения смещения необходимо создавать в базе варикапа аномальное распределение нескомпенсированных примесей с градиентом концентрации другого знака по сравнению со знаком градиента концентрации в базе диффузионного диода (рис. 3.64).

Как и в других диодах, сопротивление базы варикапа должно

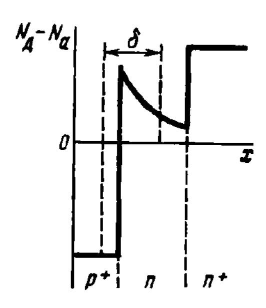


Рис. 3.64. Распределение концентрации нескомпенсированных примесей в структуре варикапа с резкой зависимостью емкости от напряжения

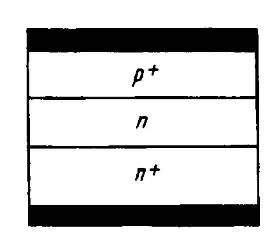


Рис. 3.65. Структура варикапа с малым сопротивлением базы

быть по возможности малым. Одновременно для большего пробивного напряжения необходимо большее удельное сопротивление слоев базы, прилегающих к *p-n*-переходу. Таким образом, база варикапа должна состоять из двух слоев (рис. 3.64 и 3.65). Основная часть базы — подложка — должна быть низкоомной. Тонкий слой базы, прилегающий к переходу, должен быть высокоомным.

Основные параметры

1. *Емкость варикапа* $C_в$ — емкость, измеренная между выводами варикапа при заданном обратном напряжении. Для различных варикапов емкость может быть от нескольких единиц до нескольких сотен пикофарад.

2. *Коэффициент перекрытия по емкости* K_C — отношение емкостей варикапа при двух заданных значениях обратных напряжений. Значение этого параметра составляет обычно несколько единиц.

3. *Добротность варикапа* $Q_в$ — отношение реактивного сопротивления варикапа на заданной частоте переменного сигнала к сопротивлению потерь при заданном значении емкости или обратного напряжения. Добротность — это величина, обратная тангенсу угла диэлектрических потерь. Добротность варикапов изме-

ряют обычно при тех же напряжениях смещения, что и емкость. Значение добротности — от нескольких десятков до нескольких сотен.

Частотные свойства

Изменение барьерной емкости при изменении напряжения на p-n-переходе и перезарядка этой емкости под действием изменяющегося напряжения обусловлены смещением основных носителей заряда в прилегающих к переходу областях. Постоянная времени этого процесса — время релаксации $\tau = \varepsilon\varepsilon_0\rho$.

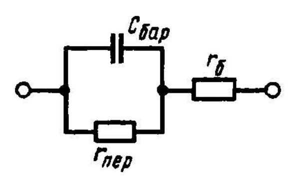


Рис. 3.66. Эквивалентная схема варикапа для широкого диапазона частот

Приняв $\rho = 10^{-1}$ Ом·см и $\varepsilon = 16$ (для германия), получаем $\tau = 1{,}4\cdot10^{-13}$ с. Значит, инерционность процессов перезарядки барьерной емкости варикапа под действием изменяющегося напряжения может сказываться только на очень высоких частотах. Однако в варикапе (как и в любом другом полупроводниковом диоде) есть объемное сопротивление базы $r_б$ и активное сопротивление p-n-перехода $r_{пер}$ (рис. 3.66), которые могут привести к снижению добротности варикапа при высоких или низких частотах.

При низких частотах в эквивалентной схеме варикапа можно не учитывать малое сопротивление базы $r_б$ по сравнению с большим емкостным сопротивлением барьерной емкости и большим активным сопротивлением перехода. Таким образом, при низких частотах упрощенная эквивалентная схема варикапа представляет собой параллельное соединение барьерной емкости и сопротивления перехода. Для параллельной схемы замещения тангенс угла диэлектрических потерь

$$\mathrm{tg}\,\delta = \frac{1}{\omega r_{пер} C_{бар}},$$

добротность

$$Q_{нч} = \omega r_{пер} C_{бар} = 2\pi f r_{пер} C_{бар}.$$

Следовательно, добротность варикапа при низких частотах растет с увеличением частоты, так как изменяется соотношение между реактивной и активной проводимостями варикапа (рис. 3.67).

Таким образом, варикапы, применяемые на очень низких частотах, должны иметь высокие значения произведения дифференциального сопротивления p-n-перехода и барьерной емкости $r_{пер}C_{бар}$. Значение этого произведения не зависит от площади перехода, так как барьерная емкость пропорциональна, а дифференциальное сопротивление обратно пропорционально площади p-n-перехода. Ясно, что дифференциальное сопротивле-

ние p-n-перехода будет больше, если значения плотности постоянных обратных токов будут меньше. Поэтому для изготовления низкочастотных варикапов целесообразно использовать полупроводниковые материалы с большой шириной запрещенной зоны (кремний, арсенид галлия и т. д.).

При высоких частотах в эквивалентной схеме варикапа можно не учитывать большое активное сопротивление перехода по сравнению с малым (при высокой частоте) параллельно включенным емкостным сопротивлением барьерной емкости. Но при этом нельзя пренебрегать сопротивлением базы, которое может оказаться сравнимым с емкостным сопротивлением барьерной емкости. Таким образом, при высоких частотах упрощенная эквивалентная схема варикапа представляет собой последовательное соединение барьерной емкости и сопротивления базы. Для последовательной схемы замещения тангенс угла диэлектрических потерь

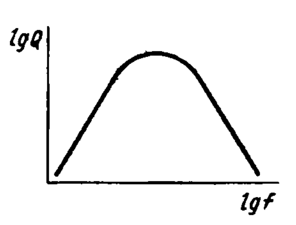


Рис. 3.67. Зависимость добротности варикапа от частоты переменного напряжения

$$\operatorname{tg}\delta = \omega r_б C_{бар},$$

добротность

$$Q_{вч} = \frac{1}{\omega r_б C_{бар}} = \frac{1}{2\pi f r_б C_{бар}}.$$

Следовательно, добротность варикапа при высоких частотах уменьшается с увеличением частоты, так как уменьшается отношение реактивного сопротивления варикапа к сопротивлению потерь (рис. 3.67).

Варикапы в основном применяют на высоких и сверхвысоких частотах. Поэтому определение одного из основных параметров варикапа — добротности — соответствует именно последовательной схеме замещения.

Таким образом, варикапы, применяемые на высоких и сверхвысоких частотах, должны иметь малые значения дифференциального сопротивления базы $r_б$. В рабочем диапазоне постоянных смещений варикапа, т. е. при обратных напряжениях, сопротивление базы можно считать неизменным. Тогда дифференциальное сопротивление базы, равное ее статическому сопротивлению, для плоскостного варикапа можно определить по формуле

$$r_б = \varrho W/S, \qquad (3.110)$$

где W — толщина базы варикапа; S — площадь p-n-перехода.

Из формулы (3.110) следует, что уменьшить дифференциальное или статическое сопротивление базы варикапа можно путем выбора исходного полупроводникового материала с большой подвижностью носителей заряда (арсенид галлия, германий с электропроводностью n-типа и т. д.).

Уменьшать сопротивление базы варикапа путем увеличения концентрации носителей заряда нецелесообразно, так как при увеличении концентрации носителей заряда будут уменьшаться подвижность носителей заряда (см. § 1.9) и пробивное напряжение варикапа (см. § 3.11).

Для уменьшения сопротивления базы без уменьшения пробивного напряжения базу варикапа делают двухслойной (см. рис. 3.65). При этом низкоомный слой базы представляет собой подложку относительно большой толщины, обеспечивающей необходимую механическую прочность структуре прибора. Высокоомный слой базы с тем же типом электропроводности имеет малую толщину (несколько микрометров). Он может быть создан методом эпитаксиального наращивания.

§ 3.32. НАДЕЖНОСТЬ ДИОДОВ

Общие понятия надежности

Надежность полупроводниковых диодов обычно значительно выше надежности других элементов радиоэлектронной аппаратуры. Однако в связи с усложнением схем требуется дальнейшее повышение надежности полупроводниковых диодов.

Количественно надежность связана с числом отказов, т.е. с числом нарушений работоспособности приборов в течение определенного промежутка времени. По характеру изменения параметров отказы полупроводниковых диодов, как и других приборов, могут быть двух видов:

1) *катастрофические,* или *внезапные,* возникающие в результате скачкообразного изменения одного или нескольких параметров прибора;

2) *условные,* или *постепенные,* возникающие в результате плавного изменения основных параметров прибора. Условность отказа определяется выбором критериев годности. Условно отказавший прибор может быть фактически работоспособным в радиоэлектронных схемах или частично утратившим свою работоспособность.

Основные причины отказов

Катастрофические отказы полупроводниковых диодов могут быть обусловлены недостатками конструкции или технологии, а также неправильной эксплуатацией диодов.

Из-за различия коэффициентов теплового расширения сочленяющихся деталей может происходить механическое нарушение контакта или растрескивание кристалла полупроводника при изменении температуры диода. Эта причина катастрофического отказа относится к конструктивным недостаткам прибора.

В высокочастотных диодах размеры электродов составляют иногда от единиц до десятков микрометров. Внутренние выводы от таких электродов выполняют в виде очень тонкой проволоки, иногда в несколько микрометров. Такую проволоку трудно изготовить строго одного сечения по длине, что может обусловить перегорание внутреннего вывода даже при нормальных нагрузках по току. К перегоранию может привести и излишняя длина внутреннего вывода, так как по длине проволоки возникают значительные перепады температур.

Наиболее частой причиной катастрофических отказов является неправильная эксплуатация диодов. Так, даже очень кратковременные импульсы токов и напряжений, превышающие допустимые значения, могут привести к необратимому пробою *р-п*-перехода диода в связи с тем, что пробой часто происходит по неоднородностям в *р-п*-переходе. При этом плотности тока в отдельных частях *р-п*-перехода могут оказаться недопустимо большими, произойдет шнурование тока с резким увеличением удельной мощности, выделяющейся в шнуре. Подобные эффекты возможны в полупроводниковых диодах даже при прохождении коротких импульсов тока с амплитудой, не превышающей допустимого значения постоянного тока.

К неправильной эксплуатации полупроводниковых диодов следует отнести размещение их вблизи нагревающихся элементов схемы, что способствует перегреву диода, а также крепление диода способом, ухудшающим условия теплообмена с окружающей средой.

Условные отказы в полупроводниковых диодах, связанные с медленным изменением их параметров, чаще всего вызваны физическими и механическими процессами на поверхности и в объеме полупроводникового кристалла, сплавов и припоев.

В § 3.14 было рассмотрено влияние поверхностных состояний на основные параметры полупроводниковых диодов. При изменении поверхностных состояний могут существенно изменяться обратные токи и пробивные напряжения диодов. Для стабилизации поверхностных состояний кристалл полупроводника с *р-п*-переходом помещают внутрь герметичного корпуса.

Одним из основных факторов, определяющих нестабильность поверхности полупроводника, является влага. Какими бы тщательными ни были герметизация прибора и предварительный отжиг деталей, все же некоторое количество влаги попадает в прибор и влияет на стабильность его характеристик. Для поглощения остаточной влаги в корпус прибора часто помещают влагопоглощающее вещество — сорбент (силикагель, алюмогель или цеолит). Однако влагопоглотитель способен выполнить свои функции только при хорошей герметичности корпуса. При наличии же микротрещин в проходных изоляторах через сотни и более часов работы диода внутрь корпуса наберется такое количество влаги, которое вызовет нестабильность электрических характеристик диода.

Микротрещины в проходных изоляторах могут возникать из-за иеудачного выбора материалов для спая стекла с коваром, иеправильиой конструкции изолятора, вследствие чего стекло работает не на сжатие, а иа растяжение, из-за нарушения необходимых допусков либо из-за небрежного обращения с прибором (изгиб внешних выводов вблизи проходных стеклянных изоляторов).

Кроме того, причиной появления условиых отказов являются физико-химические процессы в сплавах и припоях, используемых для создания *p-n*-переходов и невыпрямляющих коитактов. Иногда кристаллизация этих сплавов при изготовлении прибора идет неравномерио и образуется не предусмотренная равиовесиой диаграммой состояиия эвтектика с очень низкой температурой плавления, с ускоренными процессами диффузии различиых элементов в сплаве, что приводит к старению сплавов и ухудшению качества *p-n*-переходов и невыпрямляющих контактов.

Контрольные вопросы

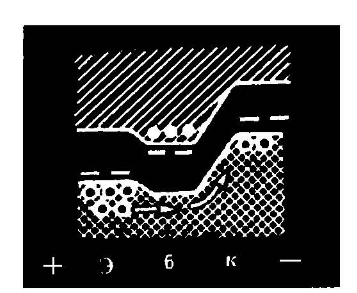
1. Что называют полупроводииковым диодом?
2. Какую область полупроводиикового диода называют базой?
3. Почему при определении плоскостных и точечных диодов в качестве характеристической длины иногда принимают диффузионную длину неосновных носителей заряда в базе диода, а иногда толщину базы?
4. Почему о плотности тока через диод можно судить по распределению неосновных носителей заряда в его базе?
5. Чем отличаются ВАХ диодов с толстой и тонкой базами?
6. Как и по каким причинам изменяется ярямая ветвь ВАХ диода с увеличением его температуры?
7. Что такое диффузионная емкость диода?
8. Как связано времи пролета неосиовных носителей заряда через базу диода с толщиной базы и коэффициентом диффузии неосиовных носителей?
9. В каких случаях процесс генерации носителей в *p-n*-переходе влияет на ВАХ диода?
10. Как влияет процесс рекомбинации носителей заряда в *p-n*-переходе диода иа его ВАХ?
11. Как связан коэффициент лавинного размножения с коэффициентом ударной ионизации?
12. Как зависит пробивное напряжение диодов при лавинном пробое от концентрации примесей в базе и от ее удельного сопротивления?
13. Как изменится пробивное иапряжение диода при лавинном и при туннельном пробое с увеличением температуры?
14. Каковы особенности теплового пробоя в реальных диодах?
15. Как могут влиять поверхностные состояния на ВАХ диода?
16. Какие явления надо учитывать при работе диода иа высоком уровие инжекции?
17. По каким причинам в базе диода возникает электрическое поле при высоком уровне инжекции?
18. Как процессы накопления и рассасывания неосновиых носителей заряда в базе диода, а также барьерная емкость влияют на работу диода при быстром изменении напряжения или тока?
19. Перечислите и объясните отличия в свойствах и параметрах кремниевых и германиевых выпрямительных диодов.
20. Что такое сопротивление растекаиия и как его можно рассчитать?

21. Какие физические явления и свойства выпрямляющих электрических переходов используются в различных полупроводниковых диодах: выпрямительных, смесительных, умножительных, модуляторных диодах и стабисторах, полупроводниковых стабилитронах и лавинно-пролетных диодах, туннельных и обращенных диодах, варикапах?

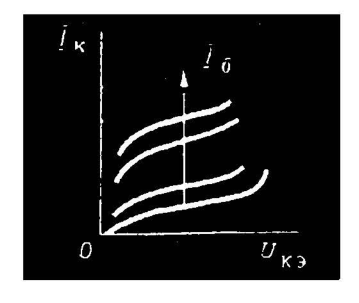
22. Почему лавинно-пролетные диоды обладают отрицательным дифференциальным сопротивлением только на определенных частотах СВЧ-диапазона?

23. Почему варикапы должны работать только при приложении к ним обратного постоянного напряжения смещения?

24. Почему в варикапах используется только барьерная емкость и не используется диффузионная емкость?



ГЛАВА 4



Биполярные транзисторы

§ 4.1. СТРУКТУРА И ОСНОВНЫЕ РЕЖИМЫ РАБОТЫ

Биполярный транзистор (обычно его называют просто транзистором) — это полупроводниковый прибор с двумя взаимодействующими выпрямляющими электрическими переходами и тремя (или более) выводами, усилительные свойства которого обусловлены явлениями инжекции и экстракции неосновных носителей заряда.

Схематическое изображение структуры биполярных транзисторов с выпрямляющими электрическими переходами в виде *p-n*-переходов приведено на рис. 4.1. Взаимодействие между *p-n*-переходами будет существовать, если толщина области между переходами (толщина базы) будет много меньше диффузионной длины неосновных носителей заряда. В этом случае носители заряда, инжектированные через один из *p-n*-переходов при его смещении в прямом направлении, могут дойти до другого перехода, находящегося под обратным напряжением, и изменить его ток. Таким образом взаимодействие выпрямляющих электрических переходов биполярного транзистора проявляется в том, что ток одного из переходов может управлять током другого перехода.

Каждый из *p-n*-переходов транзистора может быть смещен либо в прямом, либо в обратном направлении. В зависимости от этого различают три режима работы транзистора:

1) *режим отсечки* — оба *p-n*-перехода смещены в обратном направлении, при этом через транзистор проходят сравнительно небольшие токи;

2) *режим насыщения* — оба *p-n*-перехода смещены в прямом направлении, при этом через транзистор проходят относительно большие токи;

3) *активный режим* — один из *p-n*-переходов смещен в прямом направлении, а другой — в обратном направлении.

В режиме отсечки и режиме насыщения управление транзистором почти отсутствует. В активном режиме такое управление осуществляется наиболее эффективно, причем транзистор может выполнять функции активного элемента электрической схемы (усиление, генерирование, переключение и т. п.).

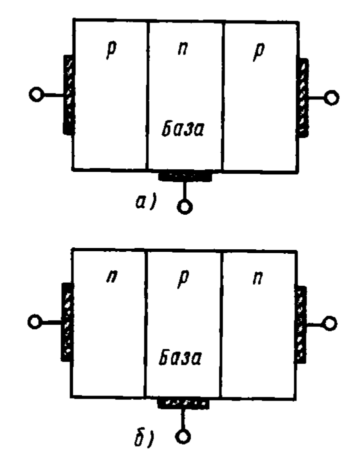


Рис. 4.1. Схематическое изображение структур биполярных транзисторов:
а — *p-n-p*-типа; *б* — *n-p-n*-типа

Область транзистора, расположенную между *p-n*-переходами, называют *базой*. Примыкающие к базе области чаще всего делают неодинаковыми. Одну из областей изготовляют так, чтобы из нее наиболее эффективно происходила инжекция носителей в базу, а другую — так, чтобы соответствующий *p-n*-переход наилучшим образом осуществлял экстракцию носителей из базы.

Область транзистора, основным назначением которой является инжекция носителей в базу, называют *эмиттером*, соответствующий *p-n*-переход — *эмиттерным*.

Область транзистора, основным назначением которой является экстракция носителей из базы, называют *коллектором*, соответствующий *p-n*-переход — *коллекторным*.

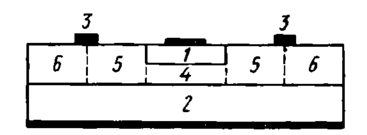


Рис. 4.2. Одна из структур реального транзистора:
1 — эмиттер; *2* — коллектор; *3* — электрод к базовой области; *4* — активная; *5* — пассивная; *6* — периферическая части базы

Если на эмиттерном переходе напряжение прямое, а на коллекторном — обратное, то включение транзистора считают *нормальным*, при противоположной полярности напряжений — *инверсным*.

Часть базы, находящуюся между эмиттером и коллектором, через которую проходят носители заряда в активном режиме работы транзистора, называют *активной частью* (рис. 4.2). Часть базы, расположенную между эмиттером и выводом базы, называют *пассивной*, а ту часть, которая лежит за выводом базы, — *периферической*.

Основные характеристики транзистора определяются в первую очередь процессами, происходящими в базе. В зависимости от распределения примесей в базе может существовать или отсутствовать электрическое поле. Если при отсутствии тока в базе существует электрическое поле, которое способствует движению неосновных носителей заряда от эмиттера к коллектору, то тран-

зистор называют *дрейфовым*, если же поле в базе отсутствует — *бездрейфовым*.

Значение напряжениости электрического поля в базе транзистора при отсутствии токов может быть определено из выражения для тока основных носителей. Например, для базы с электропроводностью n-типа

$$\mathbf{J}_n = qn_{n0}\mu_n\mathbf{E} + qD_n\text{grad}\, n_{n0} = 0; \tag{4.1}$$

отсюда

$$\mathbf{E} = -\frac{kT}{q}\frac{\text{grad}\, n_{n0}}{n_{n0}}. \tag{4.2}$$

Если учесть, что концентрация основных носителей заряда практически равно концентрации примесей, то для полупроводника n-типа получим

$$\mathbf{E} = -\frac{kT}{q}\frac{\text{grad}\, N_{д}}{N_{д}}; \tag{4.3}$$

для полупроводника p-типа

$$\mathbf{E} = \frac{kT}{q}\frac{\text{grad}\, N_{а}}{N_{а}}. \tag{4.4}$$

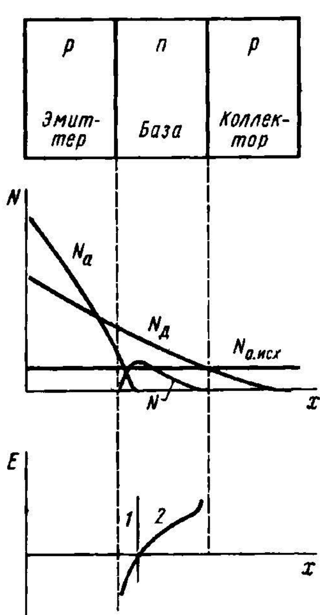


Рис. 4.3. Распределение концентрации примесей и электрического поля в базе транзистора, изготовленного методом диффузии:

1 — участок тормозящего поля; 2 — участок ускоряющего поля

Таким образом, возникновение электрического поля связано с наличием градиента концентрации примесей. Физически это объясняется тем, что градиент концентрации основных носителей заряда должен вызывать их перераспределение. При этом из области с большей концентрацией основные носители уходят, оставляя частично нескомпенсированные заряды ионов примеси, а в области с малой концентрацией примесей основные носители накапливаются. Получающееся электрическое поле всегда направлено так, что способствует движению неосновных носителей из области с большой концентрацией примесей в область с малой концентрацией. Следовательно, поле, способствующее движению неосновных носителей от эмиттера к коллектору, создается, если концентрация нескомпенсированных примесей в базе уменьшается по направлению от эмиттера к коллектору.

Такое распределение примесей в базе получается, например, при изготовлении транзистора методом диффузии (рис. 4.3). Однако в этом случае может возникнуть участок, где поле будет препятствовать движению носителей от эмиттера к коллектору (тормозящее поле), что обычно ухудшает свойства транзистора.

Различают три схемы включения транзистора: с общим эмиттером, общей базой и общим коллектором. Общим называют электрод, относительно которого измеряют и задают напряжения. Ток в цепи общего электрода не определяют. Усилительные свойства транзистора проявляются, если в схеме с общей базой в качестве входной цепи использовать эмиттерную, а в качестве выходной — коллекторную. В схеме с общим эмиттером входной является цепь базы, а выходной — цепь коллектора. Для схемы с общим коллектором входной является цепь базы, а выходной — цепь эмиттера.

Основные свойства транзистора определяются соотношениями токов и напряжений в различных его цепях и взаимным их влиянием друг на друга. Транзистор может работать на постоянном токе, малом переменном сигнале, большом переменном сигнале и импульсном сигнале.

Чтобы рассмотреть работу транзистора на постоянном токе, необходимо изучить стационарные потоки носителей в нем. Это дает возможность получить *статические характеристики и статические параметры транзистора* — соотношения между его постоянными токами и напряжениями, выраженные графически или в виде числовых значений.

На работу транзистора при малом переменном сигнале помимо стационарных потоков носителей заряда влияют также процессы накопления носителей, имеющиеся в транзисторе емкости, а также конечная скорость движения носителей. Свойства транзистора при работе на малом переменном сигнале описываются системами малосигнальных параметров.

При работе на большом сигнале и на импульсном сигнале помимо тех факторов, которые определяют работу транзистора на постоянном токе и малом переменном сигнале, существенно сказывается нелинейность статических характеристик транзистора.

§ 4.2. РАСПРЕДЕЛЕНИЕ СТАЦИОНАРНЫХ ПОТОКОВ НОСИТЕЛЕЙ ЗАРЯДА

Рассмотрим распределение потоков носителей в транзисторе на примере структуры типа *p-n-p*.

Активный режим. Через эмиттерный переход транзистора, работающего в активном режиме, происходит инжекция носителей заряда в базу (рис. 4.4, *а*). Инжектированные носители (ток $I_{\text{эр}}$) частично рекомбинируют в объеме базы и на его поверхности, а некоторые из них могут доходить до омического перехода с базой и рекомбинировать на нем (токи $I_{\text{рек}\,v}$, $I_{\text{рек}\,s}$, $I_{\text{рек.кон}}$). Остальные ин-

жектированные носители пересекают базу, доходят до коллекторного перехода и увеличивают его обратный ток.

К току носителей заряда, инжектированных эмиттером и дошедших до коллектора ($I_{Кp}$), добавляется ток носителей, образовавшихся в результате тепловой генерации в базе ($I_{Б\,ген}$), в коллекторе ($I_{К\,ген}$), а также в коллекторном переходе ($I_{ген}$). Кроме того, при достаточно больших напряжениях в коллекторном переходе

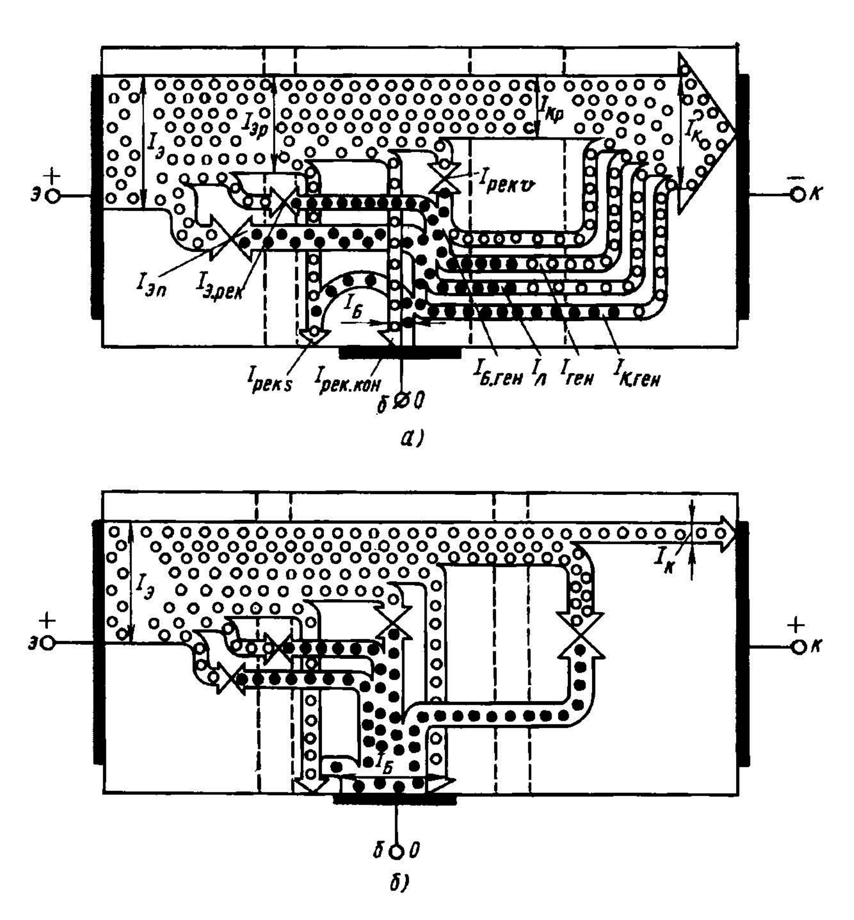


Рис. 4.4. Распределение стационарных потоков носителей заряда в транзисторе:

а — активный режим; *б* — режим насыщения

происходит лавинное размножение носителей ($I_л$). Могут существовать и токи утечки по поверхности полупроводника. Все эти токи в сумме образуют ток коллектора.

Через эмиттер помимо тока носителей заряда, инжектируемых в базу ($I_{Эp}$), проходит ток носителей, инжектируемых из базы в эмиттер ($I_{Эn}$). В области эмиттера эти носители оказываются неос-

новными и рекомбинируют. Кроме того, через эмиттерный переход проходит ток, связанный с рекомбинацией носителей в области объемного заряда ($I_{Э\,рек}$), а иногда (при малых общих токах) нужно учитывать и ток утечки.

Ток, проходящий к выводу базы, представляет собой алгебраическую сумму токов основных носителей, обусловливающих инжекцию носителей в эмиттер ($I_{Эn}$), рекомбинацию в эмиттерном переходе ($I_{Э\,рек}$) и в базе ($I_{рек\,v}+I_{рек\,s}$), обратных токов коллекторного перехода ($I_{Б\,ген}+I_{л}+I_{ген}+I_{К\,ген}$), а также тока неосновных носителей заряда, дошедших до вывода базы ($I_{рек.кон}$). Значение и направление тока базы определяются соотношением этих составляющих.

Если рассматривать распределение токов в транзисторе с точки зрения взаимного влияния эмиттера и коллектора, то оказывается, что от тока эмиттера существенно зависит только составляющая $I_{Кp}$ тока коллектора, обусловленная инжектированными эмиттером в базу носителями, дошедшими до коллектора, а также составляющая тока, связанная с лавинным размножением в коллекторном переходе ($I_{л}$). Все остальные составляющие тока коллектора либо совсем не зависят от тока эмиттера, либо эта зависимость слабая ($I_{К\,ген}$).

Кроме того, на значение тока коллектора влияет не весь эмиттерный ток, а только его составляющая $I_{Эp}$, связанная с инжекцией неосновных носителей в базу. Составляющие тока эмиттера, связанные с инжекцией из базы в эмиттер ($I_{Эn}$) и рекомбинацией в области объемного заряда ($I_{Э\,рек}$), хотя и зависят от напряжения на эмиттерном переходе, на ток коллектора непосредственно не влияют.

Чтобы ток эмиттера эффективно управлял током коллектора, принимают меры для уменьшения всех составляющих тока, не принимающих участия в таком управлении. Прежде всего стараются сократить потери носителей заряда, инжектируемых в базу. Для этого:

1) толщину базы делают небольшой по сравнению с диффузионной длиной неосновных носителей заряда, что снижает потери на рекомбинацию в объеме базы ($I_{рек\,v}$);

2) поверхность полупроводника обрабатывают так, чтобы получить по возможности малое значение скорости поверхностной рекомбинации (уменьшить $I_{рек\,s}$);

3) вывод базы располагают настолько далеко от эмиттера, что до него дырки практически не доходят (уменьшение $I_{рек.кон}$);

4) размеры коллектора делают такими, чтобы он мог перехватить весь поток носителей, идущих от эмиттера, т. е. площадь коллекторного перехода должна быть значительно больше площади эмиттерного перехода.

В результате этих мер

$$I_{Эp}\approx I_{Кp}. \qquad (4.5)$$

Для уменьшения составляющей тока носителей, инжектируе-

мых из базы в эмиттер ($I_{Эп}$), концентрацию примесей в базе делают значительно меньшей, чем в эмиттере. Это приводит к тому, что ток через эмиттерный переход состоит практически из носителей заряда, инжектируемых в базу, т. е. коэффициент инжекции (отношение дырочного тока к полному току через переход) получается высоким.

Использование для увеличения тока коллектора явления ударной ионизации чаще всего нерационально, так как приводит к сильной зависимости тока $I_к$ от напряжения на коллекторе, т. е. к нестабильной работе транзистора. Это явление используют только в лавинных транзисторах.

С учетом всех перечисленных мер можно считать, что ток коллектора транзистора в активном режиме примерно равен току эмиттера:

$$I_К \approx I_Э, \tag{4.6}$$

а ток базы значительно меньше как тока эмиттера, так и тока коллектора:

$$I_Б = I_Э - I_К. \tag{4.7}$$

Рассмотрим, как происходит усиление в транзисторе при активном режиме его работы.

В схеме с общей базой (рис. 4.5) в выходной цепи (коллекторной) практически проходит тот же ток, что и во входной (эмиттерной), т. е. усиление по току в данном случае отсутствует. Однако эта схема дает возможность получить усиление по мощности.

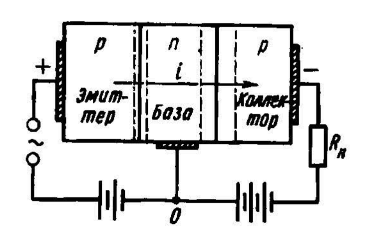


Рис. 4.5. Схема включения транзистора с общей базой

Чтобы понять принцип усиления мощности в транзисторе, да и в других усилительных приборах, надо учесть взаимодействие носителей заряда с электрическим полем. Например, дырка, двигаясь по направлению электрического поля, разгоняется в этом поле и приобретает дополнительную энергию, забирая ее от электрического поля. Если же заставить дырку двигаться против электрического поля, то она будет тормозиться этим полем, отдавая ему часть своей энергии.

Электрическое поле в коллекторном переходе транзистора состоит из постоянной составляющей, созданной внешним источником питания в цепи коллектора, и переменной составляющей, возникающей при экстракции неосновных носителей из базы в коллекторный переход. Мгновенные значения переменной составляющей электрического поля в любой момент времени направлены в сторону, противоположную постоянной составляющей.

Поэтому дырка, проходя по коллекторному переходу, взаимодействует сразу с двумя составляющими электрического поля. От постоянной составляющей электрического поля дырка забира-

ет энергию, двигаясь по направлению этой составляющей. Одновременно, двигаясь против мгновенных значений переменной составляющей электрического поля, дырка отдает часть своей энергии переменной составляющей.

Происходит своеобразное перекачивание энергии от постоянной составляющей электрического поля к переменной составляющей. Посредниками в этом перекачивании энергии являются носители заряда, инжектированные из эмиттера и дошедшие до коллекторного перехода. Для их инжекции требуется произвести относительно небольшую работу, так как высота потенциального барьера эмиттерного перехода мала.

В схеме с общим эмиттером входной цепью является цепь базы. Так как ток базы существенно меньше тока эмиттера, можно получить и усиление по току. Изменяя ток через вывод базы, меняем количество основных носителей в области базы, т. е. заряд базы, и, следовательно, потенциальный барьер между эмиттером и базой. Изменение высоты потенциального барьера вызывает соответствующую инжекцию неосновных носителей заряда. Большинство инжектированных носителей доходит до коллекторного перехода, изменяя его ток. Основной носитель заряда, введенный в базу из вывода базы, либо может исчезнуть вследствие рекомбинации, либо может быть инжектирован в эмиттер. Как указывалось, в транзисторе приняты меры, чтобы вероятность этого была мала, и на один основной носитель заряда, вошедший в базу, приходится много неосновных носителей заряда, прошедших от эмиттера до коллектора. В этом и заключается усиление по току в схеме с общим эмиттером. Усиление по мощности в данном случае объясняется аналогично усилению в схеме с общей базой.

В схеме с общим коллектором выходной цепью является эмиттерная, входной — цепь базы. В связи с тем что ток эмиттера почти равен току коллектора, здесь тоже имеет место усиление по току и по мощности.

Режим насыщения. При работе транзистора в режиме насыщения (см. рис. 4.4, *б*) в прямом направлении включен не только эмиттерный, но и коллекторный переход. Это приводит к тому, что не все носители, инжектированные эмиттером и дошедшие до коллекторного перехода, перехватываются им. Условно можно считать, что навстречу потоку неосновных носителей, идущих из базы в коллектор, идет поток таких же носителей из коллектора в базу, и суммарный их ток определяется разностью этих потоков.

В связи с тем что в режиме насыщения коллекторный переход уже не осуществляет полной экстракции носителей из базы, там происходит их накопление и интенсивная рекомбинация.

В режиме насыщения соотношения (4.5) и (4.6) не выполняются, ток базы может оказаться сравнимым с током эмиттера.

Режим отсечки. Если на обоих переходах транзистора напряжение обратное, то через них проходят токи, обусловленные

процессами тепловой генерации носителей заряда в объеме полупроводника, областях объемного заряда и на омических переходах, а также утечками. При достаточно больших напряжениях происходит лавинное размножение.

§ 4.3. РАСПРЕДЕЛЕНИЕ НОСИТЕЛЕЙ ЗАРЯДА

Значения токов в транзисторе зависят от распределения в нем неосновных носителей заряда. На это распределение влияют многие факторы — геометрические размеры транзистора, параметры его материала, состояние поверхности, токи через переходы, напряжения на них и т. д.

Активный режим. Активный режим работы транзистора характеризуется тем, что из-за наличия обратного напряже-

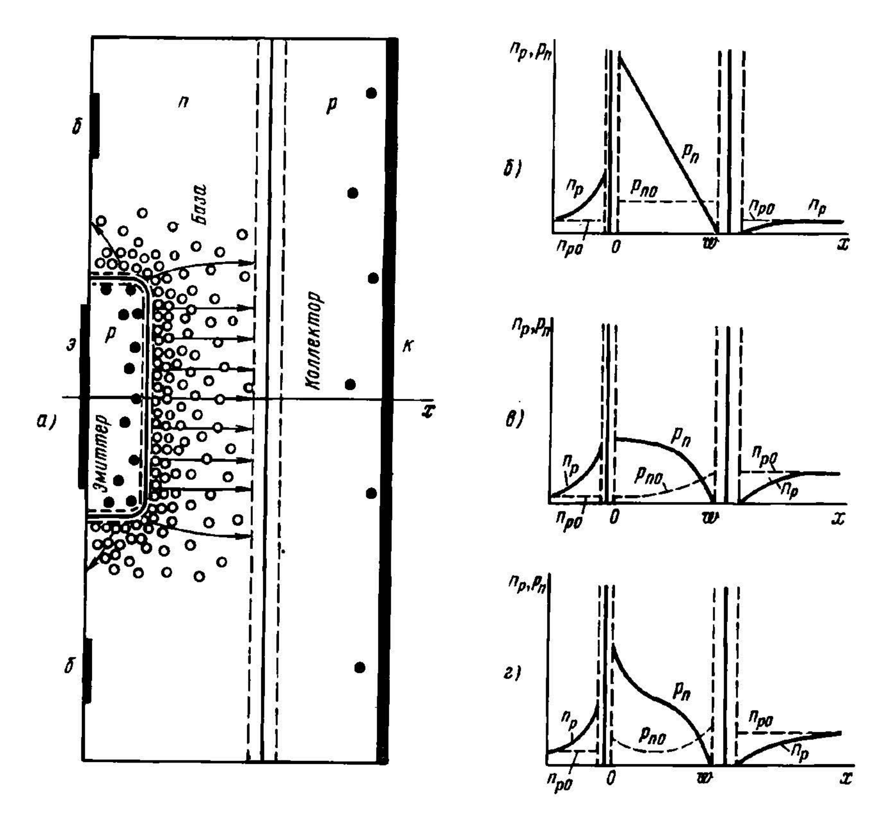


Рис. 4.6. Распределение неосновных носителей заряда в различных областях транзистора при активном режиме его работы:

а — в пространстве (схематично); *б* — в бездрейфовом транзисторе; *в* — в дрейфовом транзисторе; *г* — в транзисторе с участком тормозящего электрического поля в базе

ния коллекториый переход перехватывает практически все иосители заряда, подходящие к его граиице. В результате иосители в траизисторе движутся в иебольшой области, которая по площади примерио равна эмиттериому переходу (рис. 4.6, *а*), т. е. только через активиую часть базы. Лишь иемиогие иосители отклоияются от направлеиия, перпеидикулярного эмиттериому переходу, и некоторые из иих рекомбиикруют иа поверхиости.

В связи с тем что большииство иосителей в траизисторе, работающем в активиом режиме, движется по параллельиым путям, с достаточиой для практических целей точиостью их распределеиие можио рассматривать как одиомериое, т. е. можио считать, что коицеитрация иосителей изменяется только в направлеиии x, перпендикуляриом плоскости эмиттера или коллектора.

Закои распределеиия иосителей заряда в этом иаправлеиии определяется электрическим полем в области базы и ее толщиной w. Приближеииое аиалитическое выражеиие для такого закоиа можио получить, если учесть, что токи эмиттера и коллектора траизистора почти одииаковы. Так как иосители заряда двигаются параллельными путями, что озиачает постояиство плотиости тока неосиовиых носителей между эмиттером и коллектором, которая примерио равиа плотности тока эмиттера $J_{\text{эр}}$, уравиеиие для распределеиия, например, дырок в базе транзистора типа p-n-p примет вид

$$J_{\text{эр}} = qp_n \mu_p E - qD_p \frac{dp_n}{dx} \tag{4.8}$$

или с учетом (4.3)

$$J_{\text{эр}} = -qD_p p_n \frac{1}{N} \frac{dN}{dx} - qD_p \frac{dp_n}{dx} = $$

$$= -qD_p \frac{1}{N} \frac{d(p_n N)}{dx}, \tag{4.9}$$

где N — концентрация прнмесей в базе.

Граиичиыми условиями для случая, когда токи в траизисторе не очеиь велики, будут (рис. 4.6):
при $x=0$

$$p_n = p_{n0}(0) \exp \frac{qU_{\text{ЭБ}}}{kT}; \quad N = N(0); \tag{4.10}$$

при $x=w$

$$p_n = J_{\text{Кр}}/(qv_p); \quad N = N(w). \tag{4.11}$$

Коордниаты $x=0$ и $x=w$ соответствуют граиицам областей объемного заряда эмиттерного и коллекториого переходов (рис. 4.6).

В связи с тем что скорость движеиия иосителей заряда в области коллекториого перехода v_p довольно велика, примем, что при $x=w$ $p_n=0$.

Решив уравнение (4.9) относительно концентрации носителей в базе транзистора, получим

$$p_n = \frac{J_{эp}}{qN} \int_x^w \frac{N}{D_p} dx. \qquad (4.12)$$

Для бездрейфового транзистора $N = \text{const}$; следовательно, $D_p = \text{const}$ и

$$p_n = \frac{J_{эp}}{qD_p} (w - x). \qquad (4.13)$$

Таким образом, распределение носителей заряда, инжектированных эмиттером в базу бездрейфового транзистора, линейно (рис. 4.6, *б*). Этот вывод аналогичен полученному ранее для полупроводникового диода с тонкой базой (см. § 3.5) и, так же как для полупроводникового диода, является лишь предельным случаем. Реальное распределение носителей в базе бездрейфового транзистора отличается от линейного, хотя и незначительно.

Для транзисторов, как и для полупроводниковых диодов, можно найти распределение носителей в области эмиттера (как в диоде при прямом включении) и коллектора (как в диоде при обратном включении), что показано на рис. 4.6, *б*.

Для дрейфового транзистора при постоянном значении напряженности электрического поля в базе (это соответствует экспоненциальному распределению примесей) без учета зависимости D_p от N

$$p_n = \frac{J_{эp}}{q\mu_p E} \left\{1 - \exp\left[-\frac{qE}{kT}(w - x)\right]\right\} =$$

$$= \frac{J_{эp}}{qD_p} \left[1 - \left(\frac{N(w)}{N(0)}\right)^{1 - \frac{x}{w}}\right] \frac{w}{\ln \frac{N(0)}{N(w)}}, \qquad (4.14)$$

Соответствующее распределение концентрации носителей показано на рис. 4.6, *в*. В данном случае абсолютное значение градиента концентрации носителей растет по мере приближения к коллекторному переходу. При достаточно сильных электрических полях в базе градиент концентрации носителей вблизи эмиттера становится небольшим, т. е. ток здесь преимущественно дрейфовый. Вблизи коллектора концентрация инжектированных носителей падает, соответственно падает и дрейфовая составляющая тока, но зато растет диффузионная составляющая; поэтому полный ток остается постоянным.

При таком распределении примесей, когда создается участок тормозящего поля, распределение концентрации носителей в базе получается, как показано на рис. 4.6, *г*. Участку тормозящего электрического поля соответствует резкое возрастание концентрации носителей и ее градиента у эмиттера.

Режим насыщения. При режиме насыщения носители заряда в области базы транзистора двигаются не только от эмиттера к коллектору, но и в сторону вывода базы. Соответствеиио изменяется и их распределенне. При этом значительная концентрация неосновных носителей заряда создается и в пассивной части базы. Кроме того, прямое напряжение на коллекторном переходе приводит к инжекции в область коллектора.

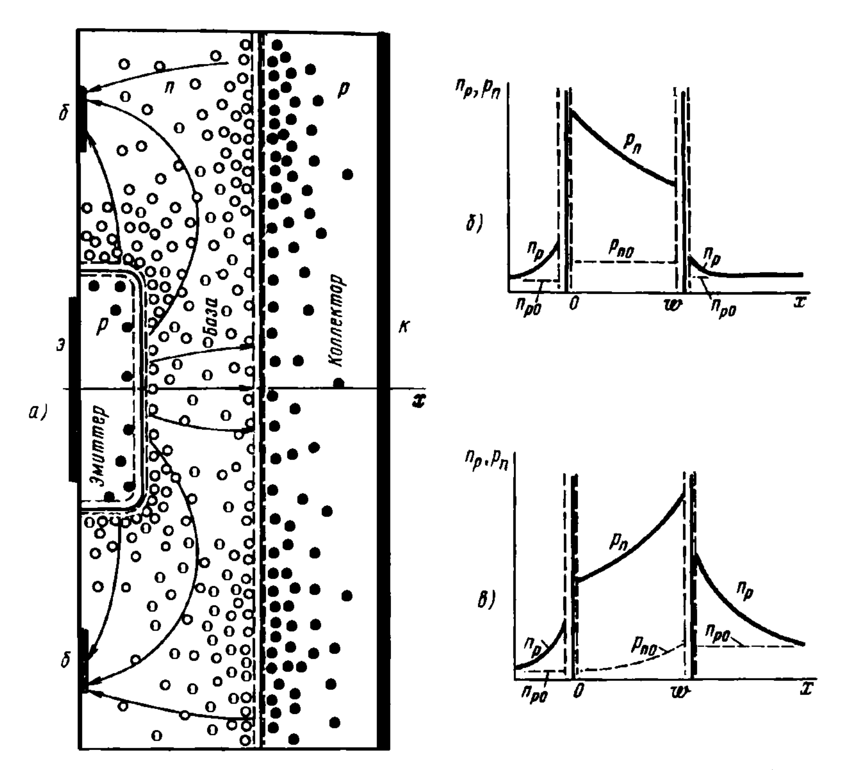


Рис. 4.7. Распределение неосновных носителей заряда в различных областях транзистора при режнме насыщения:

а — в пространстве (схематнчно); *б* — в бездрейфовом транзисторе; *в* — в дрейфовом транзисторе

Все это не дает возможности распространить на режим насыщения полученные ранее расчетным путем распределения носителей. Поэтому ограничимся качественными характеристиками (рис. 4.7). Между эмиттером и коллектором распределение носителей определяется соотношениями напряжений на *p*-*n*-переходах и напряженностью электрического поля в базе. Для бездрейфового транзнстора, если напряжение на коллекторном переходе меньше, чем на эмиттерном, концентрация неосновных носителей заряда у коллектора меньше (рис. 4.7, *б*). Из-за рекомбинации в базе закон распределения не вполне линейный.

Наличие электрического поля в базе дрейфового транзистора способствует движению носителей заряда в сторону коллектора. Кроме того, в этом случае концентрация примесей в базе у коллектора невелика; следовательно, и контактная разность потенциалов в коллекторном переходе меньше, чем в эмиттерном. Поэтому даже при меньших прямых напряжениях на коллекторе, чем на эмиттере, концентрация носителей в базе у коллекторного перехода может оказаться больше, чем у эмиттерного (рис. 4.7, *в*).

Распределение носителей в базе в направлении к ее выводу определяется инжекцией неосновных носителей через *p-n*-переходы и рекомбинацией носителей как в объеме, так и на поверхности базовой области.

Режим отсечки. При режиме отсечки части транзистора, примыкающие к его переходам, сильно обеднены неосновными носителями заряда.

§ 4.4. ПОСТОЯННЫЕ ТОКИ ПРИ АКТИВНОМ РЕЖИМЕ

Полученные в § 4.3 распределения концентрации носителей дают возможность записать аналитические выражения для постоянных токов в транзисторе, работающем в активном режиме.

Основную составляющую тока в транзисторе — ток носителей заряда, инжектированных из эмиттера в базу, находят согласно распределению носителей в базе (4.12) при подстановке $x=0$ и соответствующего значения концентрации из (4.10). Кроме того, чтобы перейти от плотности тока к току, необходимо умножить полученное выражение на площадь эмиттера $S_э$.

При расчете получим

$$I_{Эp}=\frac{S_э q n_i^2}{\int\limits_0^w \frac{N}{D_p}dx}\exp\frac{qU_{ЭБ}}{kT}. \tag{4.15}$$

Для бездрейфового транзистора

$$I_{Эp}\approx\frac{S_э q p_{n0} D_p}{w}\exp\frac{qU_{ЭБ}}{kT}, \tag{4.16}$$

что почти аналогично выражению для прямого тока полупроводникового диода с тонкой базой (см. § 3.5).

Для дрейфового транзистора при $E=\text{const}$ и $D_p=\text{const}$

$$I_{Эp}\approx\frac{S_э q p_{n0}(0) D_p}{w}\,\frac{\ln\dfrac{N(0)}{N(w)}}{1-\dfrac{N(w)}{N(0)}}\exp\frac{qU_{ЭБ}}{kT}. \tag{4.17}$$

Сравнение полученных выражений показывает, что при одинаковых размерах и напряжениях ток эмиттера дрейфового транзистора больше, чем бездрейфового.

Для определения полного тока эмиттера необходимо в соответствии с изложенным в § 4.2 добавить к $I_{Эp}$ составляющие тока, связанные с инжекцией носителей из базы в эмиттер и рекомбинацией в p-n-переходе эмиттера. Обе составляющие рассчитывают так же, как для полупроводникового диода (см. § 3.5 и 3.10):

$$I_{Эn} = \frac{S_э q n_{p0} D_p}{L_n} \left(\exp \frac{qU_{ЭБ}}{kT} - 1 \right), \tag{4.18}$$

$$I_{Эрек} \approx S_э \frac{q n_i \delta_э}{\tau} \frac{kT}{q(\varphi_{кон} - U_{ЭБ})} \exp \frac{qU_{ЭБ}}{2kT}. \tag{4.19}$$

Ток базы, обусловленный рекомбинацией носителей в ее объеме и на поверхности, можно также найти из распределения носителей. Область, где избыточная концентрация положительна, т. е. где идет процесс рекомбинации носителей, расположена между эмиттером и коллектором (рис. 4.6, *а*). Обычно концентрация инжектированных носителей заряда заметно превосходит равновесную. При этом

$$(\partial p / \partial t)_v \approx p_n / \tau_p. \tag{4.20}$$

Проинтегрировав по объему области, где создана избыточная концентрация, получим

$$I_{рек\,v} = q S_э \int_0^w \frac{p_n}{\tau_p} dx = \frac{I_{Эp}}{\tau_p} \int_0^w \left(\frac{1}{N} \int_x^w \frac{N}{D_p} dx \right) dx. \tag{4.21}$$

На поверхности базы в узкой полосе вокруг эмиттера происходнт поверхностная рекомбннация носителей заряда. В едииицу времени на единице поверхности рекомбинирует число носителей, равное $s(p_n - p_{n0})_s$ (см. § 1.14). Приняв, что вблизи эмиттера избыточная концентрация неосновных носителей иа поверхности базы

$$(p_n - p_{n0})_s \approx p_n(0) = p_{n0}(0) \exp \frac{qU_{ЭБ}}{kT}, \tag{4.22}$$

получим для тока базы, связанного с рекомбинацией на поверхности,

$$I_{рек\,s} \approx s S_{рек} \frac{\int_0^w \frac{N}{D_p} dx}{S_э N(0)} I_{Эp}. \tag{4.23}$$

Здесь $S_{рек}$ — эффективная площадь полосы по периметру эмиттера, где происходит рекомбинация инжектированных носителей. Для бездрейфовых транзисторов можно считать, что ширина этой полосы примерно равна толщине пассивной области базы.

В активном режиме неосновные носители заряда до вывода базы практически не доходят и можно считать $I_{рек.кон} = 0$.

Составляющая тока коллектора

$$I_{Кp}=I_{Эp}-(I_{рек\,v}+I_{рек\,s}). \qquad (4.24)$$

В ток коллектора входят еще составляющие, связанные с генерацией носителей в объеме коллектора и в p-n-переходе коллектора. Их определяют примерно так же, как для полупроводникового диода (см. § 3.5 и 3.10).

Ток, связанный с генерацией в коллекторе,

$$I_{К\,ген}=S_к\frac{qn_{p0}L_n}{\tau_n}. \qquad (4.25)$$

Следует учесть, что в области коллектора действует электрическое поле, связанное с прохождением тока основных носителей заряда и направленное так, что оно способствует движению неосновных носителей к коллекторному переходу. Под влиянием этого поля неосновные носители заряда могут попадать в коллекторный переход не только из областей, равных по толщине L_n, но и из более далеких.

Расстояние, которое проходит носитель заряда путем диффузии за время жизни, равно диффузионной длине. Если к диффузии добавляется дрейф, то это расстояние увеличивается до

$$L_{n\,эф}=L_n+\mu_n E\tau_n=L_n\left(1+\frac{qE}{kT}L_n\right). \qquad (4.26)$$

При этом напряженность электрического поля

$$E=\frac{J_К}{qp_{p0}\mu_p}, \qquad (4.27)$$

тогда

$$L_{n\,эф}=L_n\left(1+L_n\frac{q}{kT}\frac{J_К}{qp_{p0}\mu_p}\right) \qquad (4.28)$$

и обратный ток коллектора, связанный с генерацией носителей в его объеме,

$$I_{К\,ген}=S_э\frac{qn_{p0}L_n}{\tau_n}\left(1+L_n\frac{q}{kT}\frac{J_К}{qp_{p0}\mu_p}\right)+(S_к-S_э)\frac{qn_{p0}L_n}{\tau_n} \qquad (4.29)$$

а с учетом того, что $J_К\approx J_{Кp}$,

$$I_{К\,ген}=S_к\frac{qn_{p0}L_n}{\tau_n}+I_{Кp}\frac{n_{p0}\mu_n}{p_{p0}\mu_p} \qquad (4.30)$$

или

$$I_{К\,ген}=I_{К\,ген\,0}-(1-\bar{\alpha}^*)I_{Кp}, \qquad (4.31)$$

где

$$I_{К\,ген\,0}=S_к\frac{qn_{p0}L_n}{\tau_n};\quad \bar{\alpha}^*=1+\frac{n_{p0}\mu_n}{p_{p0}\mu_p}.$$

Значит, составляющая тока коллектора, связанная с тепловой генерацией носителей заряда, зависит от тока носителей, пришедших от эмиттера. Чтобы такое влияние было существенным, концентрации неосновных и основных носителей в коллекторе должны отличаться не очень сильно. Такие условия создаются при повышенной температуре в германиевых транзисторах, изготовленных методом диффузии.

Ток, связанный с генерацией носителей в области объемного заряда, находят обычным путем (см. § 3.10):

$$I_{\text{ген}} = \frac{q n_i \delta_{\text{к}}}{\tau} S_{\text{к}}. \qquad (4.32)$$

К току коллектора добавляется еще ток носителей заряда, генерируемых в объеме и на поверхности области базы, а также возникающих в результате лавинного размножения.

Ток генерации носителей в базе $I_{\text{Б ген}}$ можно определить так же, как обратный ток полупроводникового диода. Лавинное размножение учитывают путем умножения всех проходящих через коллекторный переход токов на коэффициент лавинного размножения M.

Тогда общий ток коллектора

$$I_{\text{К}} = M(I_{\text{К}p}\bar{\alpha}^* + I_{\text{Б ген}} + I_{\text{К ген 0}} + I_{\text{ген}}), \qquad (4.33)$$

ток базы

$$I_{\text{Б}} = I_{\text{Э}} - I_{\text{К}} = I_{\text{рек}\,v} + I_{\text{рек}\,s} + I_{\text{рек.кон}} + I_{\text{Э}n} + I_{\text{Э рек}} -$$
$$- I_{\text{К}p}(\bar{\alpha}^* M - 1) - M(I_{\text{Б ген}} + I_{\text{К ген}_0} + I_{\text{ген}}). \qquad (4.34)$$

§ 4.5. ЯВЛЕНИЯ В ТРАНЗИСТОРАХ ПРИ БОЛЬШИХ ТОКАХ

При прохождении через транзистор больших токов соотношения, выведенные в § 4.3 и 4.4, нарушаются. Это связано с рядом явлений:

1) прохождение через транзистор больших токов (точнее, токов с большой плотностью) часто сопровождается высоким уровнем инжекции. Практически в транзисторах, работающих в активном режиме, высокий уровень инжекции может наблюдаться только в базе. Действительно, эмиттер транзистора изготовляют из сильнолегированного материала, поэтому получить высокий уровень инжекции в нем довольно трудно. К тому же в общем токе эмиттера доля носителей заряда, инжектируемых из базы, невелика (см. § 4.2). В режиме насыщения высокий уровень инжекции возможен и в коллекторе;

2) прохождение больших токов через транзистор связано с перераспределением напряжений в нем из-за падений на объемных сопротивлениях;

3) при больших токах могут изменяться условия как на выпрямляющих, так и на омических переходах (последнее существенно в режиме насыщения);

4) прохождение больших токов приводит к перераспределению потенциала в областях объемного заряда *p-n*-переходов и к эффекту расширения базы.

Высокий уровень инжекции в базе. Явления при высоком уровне инжекции в базе транзистора в принципе не отличаются от рассмотренных ранее для полупроводникового диода (см. § 3.15).

В результате изменения концентрации основных носителей, связанного с условием электрической нейтральности, с одной стороны, растет инжекция носителей из базы в эмиттер (ток $I_{Эn}$), а с другой — уменьшается удельное сопротивление материала базы.

Возникающее при высоком уровне инжекции электрическое поле способствует переносу инжектированных носителей заряда через базу. При неравномерном легировании базы (когда транзистор дрейфовый) это электрическое поле накладывается на уже существующее в транзисторе. По мере повышения уровня инжекции распределение электрического поля все больше приближается к тому, которое определяется лишь инжектированными носителями, и перестает зависеть от распределения примесей.

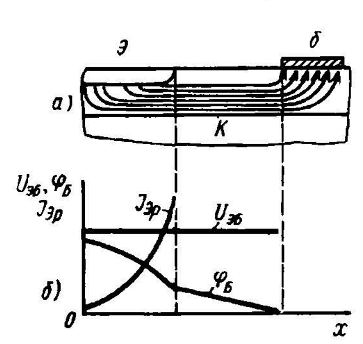


Рис. 4.8. Картина прохождения тока базы *(а)*; распределение потенциала в базе $\varphi_Б$ и плотности инжекционного тока эмиттера $J_{Эр}$ при некотором напряжении $U_{ЭБ}$ между выводами эмиттера и базы *(б)*

При высоком уровне инжекции время жизни носителей в базе также изменяется. К изменению времени жизни в объеме добавляется изменение скорости поверхностной рекомбинации. В результате соотношения между током базы и током эмиттера (4.21) и (4.23) искажаются.

Перераспределение напряжений. При прохождении токов через объемные сопротивления на них создаются соответствующие падения напряжений. В области базы наиболее существенно падение напряжения из-за прохождения тока из пространства между эмиттером и коллектором к выводу базы (рис. 4.8, *а*). Так как толщина базы очень мала, ее сопротивление может оказаться довольно большим. В результате падение напряжения на эмиттерном переходе отличается от напряжения между выводами эмиттера и базы и не одинаково (что очень существенно) в разных точках эмиттера (рис. 4.8, *б*). Это приводит к тому, что и плотность тока эмиттера в разных точках различна; причем она возрастает к краю эмиттера, т. е. по мере роста тока средние

части эмиттерного перехода как бы перестают работать — инжекция происходит преимущественно по краю эмиттера.

Изменение условий на переходах. В активном режиме работы транзистора при больших токах могут измениться условия на коллекторном переходе. Это связано (см. § 1.10) с насыщением дрейфовой скорости движения носителей. При этом концентрацию носителей в базе у коллектора нельзя считать равной нулю, как это было сделано в § 4.3, — ее значение стремится к $J_{Кp}/(qv_{pmax})$. Такое явление влечет за собой перераспределение носителей в базе. Если для бездрейфовых транзисторов из-за значительного градиента концентрации носителей во всей базе и обычно не очень больших плотностей токов влиянием такого перераспределения можно пренебречь, то для дрейфовых транзисторов оно может быть существенным.

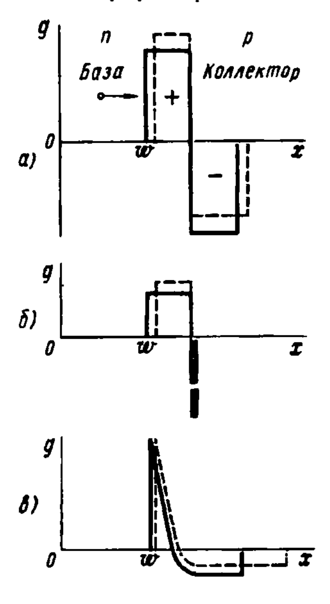


Рис. 4.9. Перераспределение плотности объемного заряда в коллекторном переходе при больших токах:

а — резкий переход; *б* — резкий переход с сильнолегированной областью коллектора; *в* — диффузионный переход

Эффект расширения базы. Перераспределение потенциала в области объемного заряда при больших токах является также следствием насыщения скорости движения носителей. Действительно, из-за того что носители заряда движутся с конечной скоростью, их концентрация в *p-n*-переходе отлична от нуля, а заряд добавляется к заряду ионизированных примесей или вычитается из него. Соответственно перераспределяется и поле в переходе.

Рассмотрим коллекторный переход транзистора (рис. 4.9, *а*). Нетрудно убедиться, что заряд, связанный с неосновными носителями, совпадает по знаку с зарядами ионизированных примесей в базе. Следовательно, из-за прохождения коллекторного тока плотность объемного заряда в части коллекторного перехода, приходящейся на базу, растет, а в части, приходящейся на коллектор, — падает. Такое изменение плотности объемного заряда эквивалентно соответствующему изменению концентрации примесей. Следовательно, при больших токах область объемного заряда в базе сужается, а в коллекторе — расширяется.

Если коллекторный переход резкий, как обычно у бездрейфовых транзисторов, и область коллектора легирована очень сильно (сплавной переход), то при больших токах изменяется практически только толщина области базы (рис. 4.9, *б*). Когда коллекторный переход плавный (дрейфовый транзистор), изменение ширины базы сопровождается сдвигом обеих границ области

объемного заряда в сторону коллектора (рис. 4.9, *в*). Если плотность тока удовлетворяет условию

$$J_{Кp} \geqslant qN_{исх}v_{p\,max},$$

то носители заряда полностью компенсируют заряд примесей в коллекторной части перехода и коллекторный переход как бы исчезает. При этом работа транзистора нарушается.

§ 4.6. СТАТИЧЕСКИЕ ПАРАМЕТРЫ

Статическими параметрами транзистора называют значения токов или напряжений, измеренных в определенных условиях, а также некоторые соотношения между этими величинами.

Параметры режима отсечки. В качестве статических параметров, характеризующих режим отсечки, обычно выбирают значения токов эмиттера и коллектора. В связи с тем что в режиме отсечки наблюдается некоторое влияние одного перехода транзистора на другой, эти токи находят при определенных условиях включения транзистора.

Начальными токами переходов называют значения токов при обратном напряжении на каком-либо переходе транзистора при условии, что напряжение на другом переходе равно нулю. Практически начальные токи находят, применяя короткое замыкание вывода соответствующей области с базой. Так, *начальный ток коллектора* $I_{КБК}$* определяют при коротком замыкании выводов эмиттера и базы, а *начальный ток эмиттера* $I_{ЭБК}$ — при коротком замыкании выводов коллектора и базы. Поскольку начальные токи переходов транзистора обычно довольно малы, короткому замыканию выводов соответствует напряжение на переходе, практически равное нулю. Согласно определению начальных токов, их находят на границе перехода транзистора из активного режима в режим отсечки.

Рассмотрим для примера происхождение тока $I_{КБК}$. При напряжении на эмиттере, равном нулю, обращаются в нуль только составляющие эмиттерного тока $I_{Эn}$ и $I_{Эрек}$, в то время как в соответствии с (4.15) при $U_{ЭБ} = 0$

$$I_{Эp} = \frac{S_э q n_i^2}{\int\limits_0^w \frac{N}{D_p} dx} \neq 0. \qquad (4.35)$$

Это объясняется тем, что при подаче напряжения на коллекторный переход в базе перераспределяются неосновные носители заряда, возникают градиент их концентрации и соответствующий

* В индексах первая буква означает вывод, в цепи которого производится измерение (Э — эмиттер, К — коллектор, Б — база), вторая буква — общий вывод, третья буква — режим измерения (К — короткое замыкание, О — разомкнутая цепь).

ток. Начальный ток коллектора складывается из тока $I_{Кp}$, также не равного нулю, и обратных токов коллекторного перехода. При не очень больших напряжениях ($M = 1$)

$$I_{КБК} = \frac{S_э q n_i^2}{\int_0^w \frac{N}{D_p} dx} - I_{рек\,v} - I_{рек\,s} - I_{рек.\,кон} + I_{Б\,ген} + I_{К\,ген} + I_{ген}. \quad (4.36)$$

Аналогично образуется и начальный ток эмиттера $I_{ЭБК}$.

Под *обратными токами переходов* транзистора понимают значения токов через какой-либо из переходов при обратном напряжении, в то время как ток в одном из двух оставшихся свободными выводов равен нулю. Так, *обратный ток коллектора транзистора в схеме с общей базой* $I_{КБО}$ определяется при токе эмиттера, равном нулю, а *обратный ток эмиттера* для той же схемы $I_{ЭБО}$ — при токе коллектора, равном нулю. Аналогично можно определить и *обратный ток коллектора для схемы с общим эмиттером* $I_{КЭО}$ (при токе базы, равном нулю).

Значение тока $I_{КБО}$ можно найти из (4.33) с учетом того, что при этом ток эмиттера отсутствует, а значит, и $I_{Кp} = 0$. Тогда

$$I_{КБО} = (I_{Б\,ген} + I_{К\,ген} + I_{ген}) M. \quad (4.37)$$

Подставив это выражение в (4.33), получим

$$I_К = M I_{Кp} + I_{КБО}. \quad (4.38)$$

Следовательно, ток $I_{КБО}$ представляет собой суммарную не управляемую со стороны эмиттера составляющую тока коллектора.

Ток базы

$$I_Б = I_{рек\,v} + I_{рек\,s} + I_{рек.кон} + I_{Эn} + I_{Э\,рек} - I_{Кp}(M - 1) - I_{КБО}. \quad (4.39)$$

Здесь тоже $I_{КБО}$ представляет собой составляющую, которая не зависит от тока эмиттера. Аналогично определяется и обратный ток эмиттера $I_{ЭБО}$.

Параметры активного режима. Статические параметры активного режима в основном характеризуют степень влияния входной цепи транзистора (эмиттерной, базовой) на выходную. К этим параметрам относится *статический коэффициент передачи тока базы*

$$h_{21Э} = \frac{I_К - I_{КБО}}{I_Б + I_{КБО}}. \quad (4.40)$$

Подставив полученные в § 4.4 выражения для токов коллектора и базы, получим

$$\frac{1}{h_{21э}} \approx \frac{1}{\overline{\alpha}^* M}\left[\frac{1}{\tau_p}\int_0^w\left(\frac{1}{N}\int_0^w\frac{N}{D_p}\,dx\right)dx + \right.$$

$$+ \frac{sS_{рек}}{S_э N(0)}\int_0^w\frac{N}{D_p}\,dx + \frac{D_n}{N(0)L_n}\int_0^w\frac{N}{D_p}\,dx + \qquad (4.41)$$

$$\left. + \frac{\delta_э}{n_i\tau}\,\frac{kT}{q(\varphi_{кон} - U_{ЭБ})}\int_0^w\frac{N}{D_p}\,dx\exp\left(-\frac{qU_{ЭБ}}{2kT}\right) - \overline{\alpha}^* M + 1\right].$$

Значение статического коэффициента передачи тока базы определяется несколькими процессами, что в формуле (4.41) отражено соответствующими сомножителями и слагаемыми.

Множитель $\overline{\alpha}^*$, учитывающий изменение обратного тока коллектора при изменении тока эмиттера, обычно примерно равен едиии-це. Коэффициент лавинного размножения M учитывает влияние этого процесса на общий ток коллектора.

Слагаемое

$$\frac{1}{\tau_p}\int_0^w\left(\frac{1}{N}\int_x^w\frac{N}{D_p}dx\right)dx = \frac{t_{прол}}{\tau_p} \qquad (4.42)$$

учитывает потери носителей заряда в базе при объемной рекомбинации.

Слагаемое

$$\frac{sS_{рек}}{S_э N(0)}\int_0^w\frac{N}{D_p}\,dx \qquad (4.43)$$

учитывает потери носителей в базе из-за поверхностной рекомбинации.

Слагаемое

$$\frac{D_n}{N(0)L_n}\int_0^w\frac{N}{D_p}dx \qquad (4.44)$$

учитывает влияние тока носителей заряда, инжектированных из базы в эмиттер.

Слагаемое

$$\frac{\delta_э}{n_i\tau}\,\frac{kT}{q(\varphi_{кон} - U_{ЭБ})}\int_0^w\frac{N}{D_p}\,dx\exp\left(-\frac{qU_{ЭБ}}{2kT}\right) \qquad (4.45)$$

учитывает потери носителей из-за рекомбинации в эмиттерном переходе.

Из перечисленных процессов только лавинное умножение и изменение обратного тока коллектора под действием тока эмиттера приводят к росту значения $h_{21э}$, остальные процессы обусловливают его снижение.

Для бездрейфового транзистора выражение (4.41) принимает более простой вид:

$$\frac{1}{h_{21э}} = \frac{1}{\bar{\alpha}^* M}\left[\frac{w^2}{2L_p^2} + \frac{sS_{рек}w}{S_э D_p} + \frac{D_n N w}{D_p N(0) L_n} + \right. \qquad (4.46)$$

$$\left. + \frac{\delta_э N w}{n_i \tau D_p}\frac{kT}{q(\varphi_{кон} - U_{ЭБ})}\exp\left(-\frac{qU_{ЭБ}}{2kT}\right) - M\left(1 + \frac{n_{p0}\mu_n}{p_{p0}\mu_p}\right) + 1\right].$$

Коэффициент передачи тока базы — один из важнейших параметров транзистора, поэтому целесообразно рассмотреть влияние на него значения постоянной составляющей тока.

При малых напряжениях, как известно (см. § 3.10), в образовании тока через p-n-переход большую роль играет рекомбинация в p-n-переходе. Так как рекомбинационный ток не связан с инжекцией в базу, он приводит к снижению $h_{21э}$. В формуле (4.41) это отражено в том, что слагаемое, учитывающее рекомбинацию в p-n-переходе (4.45), растет с уменьшением $U_{ЭБ}$.

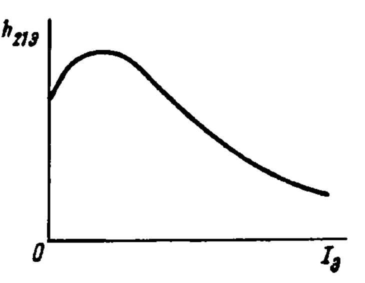


Рис. 4.10. Зависимость коэффициента передачи постоянного тока базы от постоянного тока эмиттера

По мере роста напряжения на эмиттере и тока эмиттера роль рекомбинации в p-n-переходе уменьшается, соответственно растет $h_{21э}$. К тому же с увеличением тока повышается уровень инжекции, а возникающее при этом поле способствует движению инжектированных носителей заряда через базу. Действие такого поля эквивалентно росту D_p (см. § 3.15), что приводит к уменьшению первого и второго слагаемых в формуле (4.41).

Одновременно при повышении уровня инжекции начинает расти концентрация носителей в базе, что ведет к росту тока носителей, инжектированных из базы в эмиттер. В формуле (4.46) вместо N в третьем слагаемом следовало бы взять $N + p_n$. Соответственно уменьшится $h_{21э}$.

К снижению $h_{21э}$ при больших токах приводит и вытеснение носителей к краю эмиттера, рассмотренное в § 4.5, в связи с чем возрастает роль поверхностной рекомбинации.

В некоторых транзисторах уменьшение $h_{21э}$ при больших токах может быть связано с увеличением толщины базы. На зависимость от тока влияет также изменение времени жизни носителей в объеме и скорости поверхностной рекомбинации. Результирующая зависимость $h_{21э}$ от $I_э$ чаще всего имеет вид кривой с максимумом (рис. 4.10). Обычно значение $h_{21э}$ составляет несколько десятков.

Иногда вводят *статический коэффициент передачи тока эмиттера*

$$h_{21Б} = \frac{I_К - I_{КБО}}{I_Э}, \tag{4.47}$$

при этом

$$h_{21Э} \approx \frac{|h_{21Б}|}{1 - |h_{21Б}|}. \tag{4.48}$$

Абсолютное значение $h_{21Б}$, как правило, лишь немного меньше единицы.

Помимо этого иногда используют значение крутизны передаточной характеристики

$$S = \frac{I_К - I_{КБО}}{U_{ЭБ}}. \tag{4.49}$$

Формулы для расчета этих параметров можно вывести примерно так же, как для $h_{21Э}$.

Используя статические параметры и выражение для тока $I_{КБО}$, перепишем выражение для тока $I_{КБК}$:

$$I_{КБК} = h_{21Б} \frac{q n_i^2 S_э}{\int\limits_0^w \frac{N}{D_p} dx} + I_{КБО}. \tag{4.50}$$

Параметры режима насыщения. В качестве параметров режима насыщения обычно принимают измеренные при определенных токах значения падения напряжения между выводами транзистора. Например, $U_{КБ\,нас}$ — падение напряжения между выводами коллектора и базы в режиме насыщения (для схемы с общей базой) или $U_{КЭ\,нас}$ — падение напряжения между выводами коллектора и эмиттера в режиме насыщения (для схемы с общим эмиттером).

Если эти падения напряжения отнести к току, проходящему через коллектор, то полученный параметр называют *сопротивлением насыщения.*

§ 4.7. ПРОБОЙ ТРАНЗИСТОРОВ

На процессы в транзисторе существенное влияние оказывает напряжение на коллекторе. Такое влияние обусловлено следующим. При изменении напряжения изменяется толщина области объемного заряда коллекторного перехода и соответственно толщина базы, а при достаточно больших коллекторных напряжениях начинает сказываться лавинное размножение.

С повышением напряжения на коллекторе толщина базы становится меньше, что приводит к увеличению градиента концентрации носителей заряда в базе, к уменьшению времени, в течение которого носители находятся в базе и, следовательно, к умень-

шению роли рекомбинации в базе. Это ведет к росту коэффициентов передачи $h_{21Э}$ и $h_{21Б}$.

Смыкание переходов. При достаточно больших напряжениях на коллекторном переходе область объемного заряда коллекторного перехода может достигнуть эмиттерного перехода — произойдет так называемое смыкание переходов. При этом потенциальный барьер эмиттерного перехода понижается (рис. 4.11), возрастает ток эмиттера, а значит, и ток коллектора. По внешним признакам смыкание напоминает пробой или короткое замыкание эмиттера с коллектором. Таким образом, смыкание переходов является одной из причин, ограничивающих напряжение коллектора.

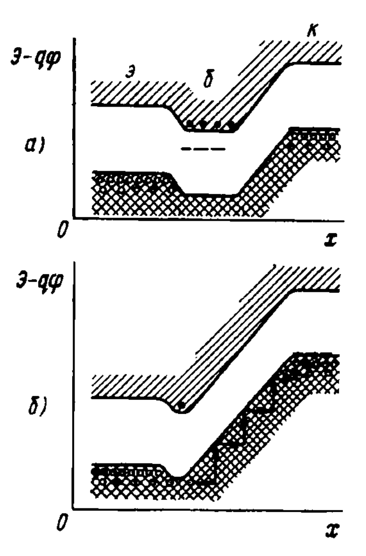


Рис. 4.11. Эффект смыкания переходов в транзисторе:

а — энергетическая диаграмма транзистора при небольшом напряжении на коллекторе; *б* — смыкание областей объемного заряда коллекторного и эмиттерного переходов

Лавинный пробой транзистора в схеме с общей базой. Второй причиной, ограничивающей напряжение коллектора, является лавинное размножение. При этом существенную роль играет режим цепи базы. Если ток в цепи базы не ограничен, что имеет место, например, в схеме с общей базой, то пробой транзисторов не отличается от пробоя полупроводникового диода. В этом случае в коллекторном переходе произойдет лавинный пробой при пробивном напряжении $U_{КБО\,проб}$.

Лавинный пробой коллекторного перехода представляет собой обратимый процесс, если ограничить возрастающий при пробое ток. С увеличением тока коллектора при лавинном размножении лавинный пробой может перейти в тепловой пробой с появлением отрицательного дифференциального сопротивления на выходе транзистора. Этот переход к тепловому пробою наиболее вероятен в транзисторах, изготовленных из германия (материала с малой шириной запрещенной зоны).

Пробой транзистора в схеме с общим эмиттером. С увеличением напряжения на коллекторном переходе статический коэффициент передачи тока эмиттера с учетом размножения носителей в коллекторном переходе возрастает и при напряжении $U_{КЭ\,проб}$ становится равным единице. Это напряжение при некоторых условиях может оказаться пробивным для транзистора в схеме с общим эмиттером. Действительно, ток коллектора в схеме с общим эмиттером

$$I_К = \frac{I_{КБО}}{1-|h_{21Б}|} + \frac{|h_{21Б}|}{1-|h_{21Б}|} I_Б. \qquad (4.51)$$

При неизменном токе базы ($I_Б = \text{const}$) и при $U_{КЭ} \rightarrow U_{КЭ\,проб}$ ток коллектора будет стремиться к бесконечности, т. е. произойдет пробой транзистора, включенного по схеме с общим эмиттером.

Какие же физические процессы приводят к пробою транзистора при данных условиях? Если ток базы зафиксирован (например, при разомкнутой цепи базы или при включении в нее достаточно большого сопротивления), то в транзисторе начинает проявляться обратная связь. Образующиеся при лавинном размножении пары носителей заряда разделяются электрическим полем коллекторного перехода: неосновные для базы носители уходят в коллектор, а основные — в базу (рис. 4.12). Таким образом, в базе создается избыточный заряд основных носителей и соответственно изменяется ее потенциал. Получающееся при этом напряжение открывает эмиттерный переход и увеличивает ток эмиттера.

Если вывод базы отключен от схемы, т. е. ток базы равен нулю, то основные носители заряда, накопившиеся в базе, могут исчезнуть только двумя путями: либо уйти в эмиттер, либо рекомбинировать с неосновными носителями, инжектированными эмиттером. Однако транзистор делают так, что вероятность этих событий довольно мала — из эмиттера в базу проходит гораздо больше носителей, чем из базы в эмиттер, и носители, инжектированные в базу, почти не рекомбинируя, доходят до коллектора. Следовательно, каждый основной носитель, оказавшийся в базе в результате ударной ионизации в коллекторном переходе, при отмеченных условиях вызовет инжекцию из эмиттера в базу большого числа неосновных носителей, что и приведет к существенному росту тока коллектора.

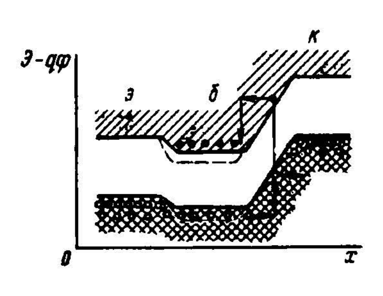


Рис. 4.12. Изменение потенциальных барьеров в транзисторе при лавинном размножении носителей заряда в коллекторном переходе и токе базы, равном нулю

Пробивное напряжение коллектор — эмиттер при $I_Б = 0$ может быть значительно меньше пробивного напряжения коллектор — база при $I_Э = 0$.

Соотношение между этими пробивными напряжениями транзистора в различных схемах включения можно найти, используя выражение для коэффициента лавинного размножения носителей (3.70):

$$|h_{21Б}| = |h_{21Б}|_{M=1} M = \frac{|h_{21Б}|_{M=1}}{1 - \left(\frac{U_{КЭО\,проб}}{U_{КБО\,проб}}\right)^{b}} = 1. \quad (4.52)$$

Тогда

$$U_{\text{КЭОпроб}} = U_{\text{КБОпроб}} \sqrt[b]{1 - |h_{21\text{Б}}|_{M=1}}, \qquad (4.53)$$

где $|h_{21\text{Б}}|_{M=1}$ — абсолютиое значение $h_{21\text{Б}}$ без учета лавииного размиожеиия.

Согласио выражеиию (4.53), разиица между пробивными иапряжеииями траизистора при отключеииом третьем выводе тем больше, чем больше статический коэффициеит передачи тока эмиттера и чем меиьше коэффициент b, который зависит от исходного материала (см. § 3.11).

Отсюда можно сделать следующие практические выводы:

1. Необходимо иметь в виду возможиость пробоя транзистора, включениого по схеме с общим эмиттером, при зиачительно меиьших иапряжеииях, чем пробивное иапряжеиие коллекторного перехода, если в цепь базы включено отиосительно большое сопротивление. Эти процессы будут обратимыми, если ток коллектора ограиичеи параметрами виешией схемы в цепи коллектора (например, при большом сопротивлении иагрузки). В противиом случае мощиость, выделяющаяся в коллекториом переходе при резком увеличеиии тока коллектора, может превысить допустимое зиачеиие, тогда произойдет иеобратимый пробой траизистора, связаииый с тепловым разрушеиием.

2. При включеиии траизистора в схему, иаходящуюся под иапряжеиием (например, при измерении параметров транзисторов), виачале необходимо присоедииить вывод базы, а затем выводы эмиттера и коллектора, чтобы не возникло условие иулевого тока базы.

Вторичный пробой. Под вторичным пробоем поимают явлеиия, связаииые с разогревом коллекториого перехода и приводящие к резкому увеличению коллекториого тока при одновремеииом умеиьшеиии коллекториого иапряжения. При вторичном пробое траизистора, как и при тепловом пробое диода, происходит шиурование тока, проходящего через коллекториый переход.

Шиурование тока связаио с иаличием различиого рода дефектов на поверхиости и в объеме траизисториой структуры, которые могут приводить к локальиому увеличеиию плотиости тока через коллекториый переход. Локальиое увеличение плотиости тока приводит к локальиому разогреву, что, в свою очередь, обусловливает:

1) увеличеиие тепловой генерации иосителей заряда в этом месте p-n-перехода коллектора и, следовательио, увеличеиие локальиой плотиости тока, еще больший локальиый разогрев и т. д. При этом тепловая геиерация носителей может возрасти иастолько, что область объемиого заряда вообще исчезиет иа малой площади коллекторного перехода;

2) локальное уменьшение толщины p-n-перехода коллектора и повышеиие напряженности поля, а следовательио, и увеличение ударной ионизации. При этом увеличатся локальная плотиость тока, локальиый перегрев, тепловая генерация и т. д.;

3) увеличение локального коэффициента передачи тока эмиттера, так как с повышением температуры растет время жизни носителей; это вызывает еще большую концентрацию проходящего тока, еще больший локальный разогрев и т. д.

Все эти явления приводят к резкому увеличению тока и уменьшению напряжения. Инерционность данных явлений, связанных с тепловыми процессами, может быть очень малой из-за малости объема, где происходит шнурование тока.

Если ток через транзистор при вторичном пробое не ограничить, то локальный разогрев приведет к необратимым изменениям в транзисторе вплоть до локального расплавления. Чаще всего в результате вторичного пробоя базовая область транзистора проплавляется эмиттерным сплавом, т. е. сквозь базу протягивается тонкая область того же типа электропроводности, что и эмиттерная и коллекторная области, закорачивая эмиттер и коллектор. При этом характеристики как эмиттерного, так и коллекторного переходов могут остаться неизменными, а эмиттерная и коллекторная области окажутся закороченными.

Если же ток через транзистор при вторичном пробое ограничить, то локальный разогрев может не привести к расплавлению кристалла в тонкой области шнурования тока. В этом случае иногда может существовать относительно стабильное состояние с малым падением напряжения между коллектором и эмиттером. Однако сохранение подобного состояния в течение длительного времени или неоднократные повторения вторичного пробоя обычно приводят к необратимым изменениям параметров транзистора в связи с большой локализацией выделяющейся мощности.

Необходимо отметить, что отсутствие дефектов в структуре транзистора не гарантирует от возникновения вторичного пробоя. Так, базовая область под *p-n*-переходом может быть не эквипотенциальна из-за прохождения базовых токов. При различных направлениях тока базы наблюдается увеличение плотности тока эмиттера либо по периферии эмиттерного перехода, либо в центре перехода. Этот эффект может также создавать предпосылки для развития процессов шнурования тока.

Вероятность вторичного пробоя возрастает с увеличением рабочего тока транзистора и напряжения на коллекторе, так как при этом неравномерность выделения мощности в транзисторе проявляется сильнее из-за большего перегрева и, следовательно, из-за большего влияния собственной электропроводности.

§ 4.8. СТАТИЧЕСКИЕ ХАРАКТЕРИСТИКИ

Системы статических характеристик

Если обозначить напряжение и ток входного электрода транзистора через U_1 и I_1, а напряжение и ток выходного — через U_2 и I_2, то взаимозависимость этих четырех величин можно выразить

двадцатью четырьмя семействами характеристик, отиосящихся к шести системам:

1) $U_1 = f_1(I_1,\ I_2),$
$U_2 = f_2(I_1,\ I_2);$

2) $I_1 = f_1(U_1,\ U_2),$
$I_2 = f_2(U_1,\ U_2);$

3) $U_1 = f_1(I_1,\ U_2),$
$I_2 = f_2(I_1,\ U_2);$

4) $I_1 = f_1(U_1,\ I_2),$
$U_2 = f_2(U_1,\ I_2);$

5) $I_1 = f_1(I_2,\ U_2),$
$U_1 = f_2(I_2,\ U_2);$

6) $I_2 = f_1(U_1,\ I_1),$
$U_2 = f_2(U_1,\ I_1).$

Из четырех возможных семейств характеристик каждой системы, которые связывают между собой четыре величины: I_1, I_2, U_1 и U_2, два семейства являются основными, а два — второстепеииыми. Их можно получить из осиовных путем перестроеиия. В качестве основных удобно выбирать семейства характеристик, связывающих ток и напряжение на входе — *входные характеристики* — и ток и напряжение на выходе — *выходные характеристики*. Другие два семейства характеристик являются следствием входных и выходных. Семейства характеристик, которые связывают токи или напряжения на выходе с токами или напряжеииями на входе, называют *характеристиками передачи,* а семейства, которые связывают напряжения или токи на входе с токами или напряжениями на выходе, — *характеристиками обратной связи.*

Из шести систем статических характеристик как наиболее удобные применяют первые три, но иаиболее широкое распространеиие получила система

$$U_1,\ I_2 = f(I_1,\ U_2). \tag{4.54}$$

Важиым преимуществом такой системы является то, что при исследовании статических характеристик в качестве независимых переменных берут ток на входе и напряжение на выходе транзистора. В схемах с общим эмиттером и с общей базой входной ток проходит через открытый эмиттерный переход, сопротивление которого мало, так что поддерживать заданное зиачение тока на входе удобно, включив последовательно относительно большое сопротивление. Необходимые иапряжения на выходе поддерживаются постоянными, если выходную цепь питать от источиика напряжения. На практике это осуществить легко, поскольку выходная цепь транзистора обладает высоким сопротивлением.

Кроме того, система уравнений (4.54) соответствует наиболее распространенной системе малосигнальных параметров (h-параметров), которая, в свою очередь, имеет ряд преимуществ перед другими системами малосигнальных параметров (см. § 4.10).

Характеристики транзистора, включенного по схеме с общей базой

Входные характеристики. Семейство входных статических характеристик транзистора, включенного по схеме с общей базой, представлено на рис. 4.13, *а*.

Общий характер этих зависимостей определяется *p-n*-переходом эмиттера, включенным в прямом направлении. Поэтому по внешнему виду входные характеристики похожи на прямые ветви ВАХ диода.

Смещение входных статических характеристик вниз в выбранной системе координат при увеличении абсолютного значения напряжения на коллекторе объясняется тем, что напряжение на коллекторе влияет на концентрацию носителей заряда около него

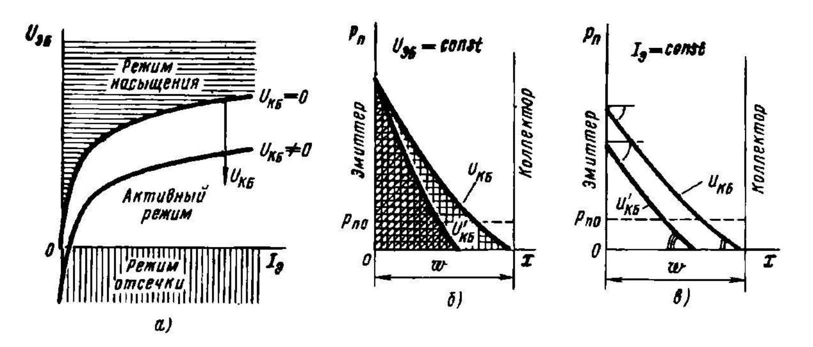


Рис. 4.13. Входные статические характеристики транзистора, включенного по схеме с общей базой (*а*); распределение концентрации дырок в области базы при постоянном напряжении на эмиттерном переходе (*б*) и при постоянном токе эмиттера (*в*); $|U'_{КБ}|>|U_{КБ}|$

и изменяет толщину базы из-за изменения толщины коллекторного перехода. Эти причины приводят к увеличению градиента концентрации неосновных носителей (дырок) в базе с увеличением абсолютного значения напряжения на коллекторе при постоянном напряжении на эмиттере (рис. 4.13, *б*).

Следовательно, с увеличением абсолютного значения напряжения на коллекторе и при постоянном напряжении на эмиттере возрастает ток эмиттера, т. е. характеристики смещаются вниз.

Тот же результат можно получить, если рассмотреть уменьшение концентрации неосновных носителей заряда в базе транзистора около *p-n*-перехода эмиттера при увеличении абсолютного значения напряжения на коллектора и при постоянном токе эмиттера (рис. 4.13, *в*).

Через эмиттерный *p-n*-переход проходит ток и при напряжении на эмиттере, равном нулю. Чтобы ток эмиттера стал равным нулю, на эмиттер должно быть подано обратное напряжение (работа в режиме отсечки).

Выходные характеристики. Семейство выходных статических характеристик транзистора, включенного по схеме с общей базой, представлено на рис. 4.14.

Общий характер этих зависимостей аналогичен обратной ветви ВАХ диода, так как коллекторный переход включен в обратном направлении.

Смещение выходных статических характеристик вверх в выбранной системе координат при увеличении тока эмиттера соответствует принципу действия транзистора.

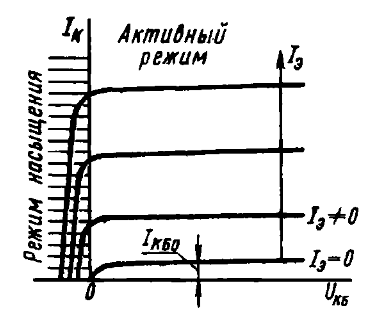


Рис. 4.14. Выходные статические характеристики транзистора, включенного по схеме с общей базой

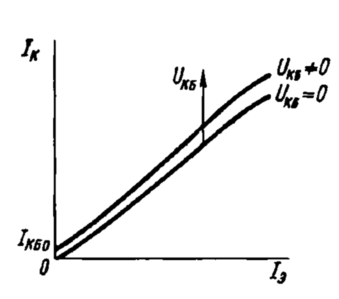


Рис. 4.15. Статические характеристики передачи тока транзистора, включенного по схеме с общей базой

Через коллекторный переход проходит ток и при напряжении на коллекторе, равном нулю, т. е. при коротком замыкании коллектора с базой. Это связано с наличием градиента концентрации неосновных носителей заряда в базе транзистора при инжекции носителей из эмиттера, т. е. при существовании тока эмиттера. Чтобы ток коллектора стал равным нулю, на коллектор должно быть подано прямое напряжение (работа в режиме насыщения).

Характеристики передачи тока. Семейство статических характеристик передачи тока транзистора, включенного по схеме с общей базой, показано на рис. 4.15.

Общий характер этих зависимостей свидетельствует о том, что в транзисторе $I_К = I_{КБО} + |h_{21Б}|I_Э$. Характеристики передачи в первом приближении можно считать прямыми линиями. В действительности коэффициент передачи постоянного тока эмиттера зависит от тока эмиттера (причины зависимости рассмотрены в § 4.5 и 4.6).

Смещение статических характеристик передачи вверх в выбранной системе координат при увеличении напряжения на коллекторе связано с уменьшением рекомбинации дырок при их прохождении через более тонкую базу (рис. 4.13, *в*).

При напряжении на коллекторе, отличном от нуля, характе-

ристики передачи выходят не из начала координат, а из точек на оси ординат, соответствующих обратному току коллектора $I_{КБО}$. Однако значением этого тока часто можно пренебречь.

Характеристики обратной связи. Семейство статических характеристик обратной связи транзистора, включенного по схеме с общей базой, приведено на рис. 4.16.

Общий характер этих зависимостей можно выяснить путем анализа рис. 4.13, *в*. Характеристики обратной связи могут быть легко получены из семейства входных статических характеристик путем графического перестроения.

Смещение статических характеристик обратной связи вверх при увеличении тока эмиттера очевидно и не требует специального пояснения.

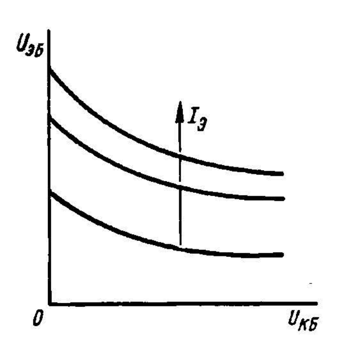


Рис. 4.16. Статические характеристики обратной связи транзистора, включенного по схеме с общей базой

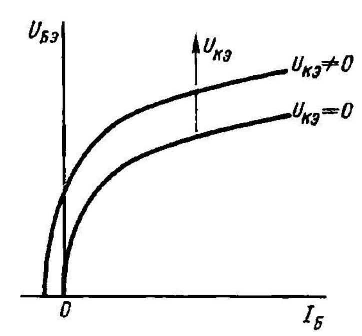


Рис. 4.17. Входные статические характеристики транзистора, включенного по схеме с общим эмиттером

Характеристики транзистора, включенного по схеме с общим эмиттером

Входные характеристики. Семейство входных статических характеристик транзистора, включенного по схеме с общим эмиттером, представлено на рис. 4.17.

Общий характер этих зависимостей аналогичен характеру подобных характеристик транзистора, включенного по схеме с общей базой, так как ток базы является суммой обратного тока коллектора $I_{КБО}$ и рекомбинационной составляющей, которая примерно пропорциональна току эмиттера и представляет собой малую часть тока эмиттера.

Смещение входных статических характеристик вверх в выбранной системе координат при увеличении абсолютного значения напряжения на коллекторе связано с уменьшением общего количества неосновных носителей заряда в базе (заштрихованные площади на рис. 4.13, *б*) и, следовательно, с уменьшением коли-

чества рекомбинирующих носителей. Поэтому уменьшается составляющая тока базы, обусловленная рекомбинацией, при постоянном напряжении между базой и эмиттером.

При наличии напряжения на коллекторе и токе базы, равном нулю, на выводе базы существует напряжение, которое объясняется падением напряжения на сопротивлении *p-n*-перехода эмиттера. Поэтому входные статические характеристики для схемы с общим эмиттером заходят во второй квадрант при $U_{КЭ} \neq 0$.

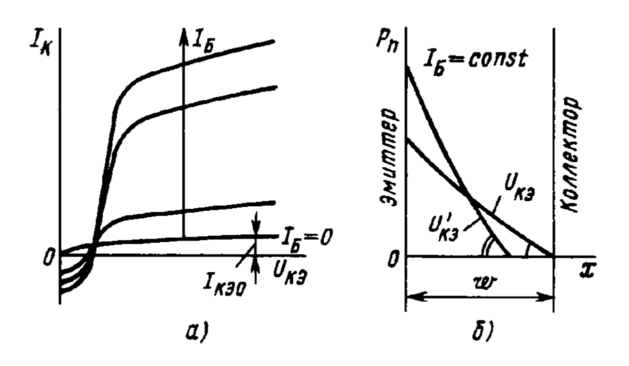


Рис. 4.18. Выходные статические характеристики транзистора, включенного по схеме с общим эмиттером (*а*), и распределение концентрации дырок в области базы при постоянном токе базы (*б*); $|U'_{КЭ}|>|U_{КЭ}|$

Выходные характеристики. Семейство выходных статических характеристик транзистора, включенного по схеме с общим эмиттером, дано на рис. 4.18, *а*.

Общий характер этих зависимостей аналогичен характеру обратной ветви ВАХ диода, так как бо́льшая часть напряжения источника питания выходной цепи падает на *p-n*-переходе коллектора, включенном в обратном направлении. Однако в отличие от выходных характеристик схемы с общей базой выходные характеристики схемы с общим эмиттером имеют значительно больший наклон, т. е. наблюдается бо́льшая зависимость тока коллектора от напряжения на коллекторе. Причины этого явления поясняет рис. 4.18, *б*, на котором видно, что с увеличением напряжения на коллекторе при постоянном токе базы увеличивается ток эмиттера и соответственно растет ток коллектора. Напомним, что основная составляющая тока базы (рекомбинационная) приблизительно пропорциональна общему числу дырок в базе и, следовательно, пропорциональна площади под кривыми распределения концентрации дырок в базе.

Смещение выходных статических характеристик вверх происходит в соответствии с формулой (4.40) и связано с увеличением тока эмиттера при условии постоянства напряжения на коллекторе и увеличения тока базы. Необходимо отметить различное рас-

стояние между выходными характеристиками по оси токов при равных приращениях тока базы. При малых токах базы кривые располагаются часто, при больших токах базы — редко, а затем опять часто. Неравномерность расположения характеристик связана с изменением коэффициента передачи постоянного тока базы ($h_{21Э}$) при изменении тока (см. рис. 4.10).

В транзисторе, включенном по схеме с общим эмиттером, ток $I_{КЭО}$ превосходит ток $I_{КБО}$ для схемы с общей базой ($I_{КЭО} = \frac{I_{КБО}}{1 - |h_{21Б}|}$. Объясняется это тем, что при токе базы, равном нулю, и при подаче напряжения на коллектор относительно эмиттера (в схеме с общим эмиттером) *p-n*-переход эмиттера оказывается включенным под небольшое прямое напряжение. Поэтому обратный ток коллектора из-за инжектированных из эмиттера в базу дырок увеличивается.

При напряжении на коллекторе, равном нулю, т. е. при коротком замыкании коллектора с эмиттером, и при наличии тока базы *p-n*-переход коллектора оказывается включенным в прямом направлении, так как он, по существу, включен параллельно *p-n*-переходу эмиттера. При $I_К = 0$ и $I_Б \neq 0$ из эмиттера происходит инжекция дырок, что обеспечивает около коллектора в базе их концентрацию, превосходящую равновесное значение. Если же концентрация неосновных носителей заряда в базе около *p-n*-перехода превышает равновесную, то это соответствует прямому включению перехода. Таким образом, транзистор работает в режиме насыщения при напряжении на коллекторе, равном нулю, и даже при небольшом запирающем напряжении на коллекторе относительно эмиттера.

Значение этого напряжения определяет сопротивление насыщения транзистора (см. § 4.6). На это напряжение влияют соотношение падений напряжения на эмиттерном и коллекторном переходах, объемное сопротивление коллектора и сопротивление базы.

На объемном сопротивлении коллектора при прохождении тока создается падение напряжения, направленное так, что оно открывает коллекторный переход. Поэтому напряжение на внешнем выводе коллектора, соответствующее выходу транзистора из режима насыщения, возрастает.

Токи в базе транзистора направлены так, что создаваемое ими падение напряжения запирает наиболее удаленные от вывода базы части эмиттерного и коллекторного переходов (см. § 4.5). Вследствие этого даже тогда, когда средняя часть коллекторного перехода уже запирается, его периферические области еще остаются открытыми и через них проходит ток к выводу базы (см. рис. 4.7, *а*). В результате ток эмиттера замыкается не через внешний вывод коллектора, а через область коллектора и вывод базы. Такое явление также приводит к росту напряжения, при котором транзистор выходит из режима насыщения. Диапазон этих значений напряжения коллектора тем больше, чем больше

ток базы. Соответственно выходные статические характеристики транзистора в схеме с общим эмиттером при малых напряжениях на коллекторе и при $I_Б \neq 0$ заходят в четвертый квадрант.

При изменении направления входного (базового) тока, т. е. при изменении полярности постоянного напряжения на базе относительно эмиттера, уменьшается значение тока эмиттера, что приводит к уменьшению коэффициента передачи тока эмиттера. Следовательно, коэффициент передачи тока эмиттера достигнет единицы при бо́льших значениях напряжения на коллекторе,

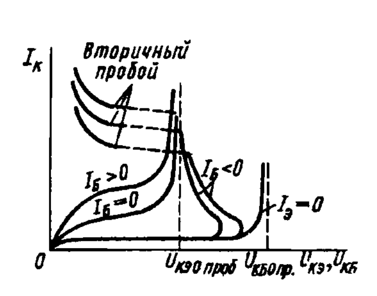


Рис. 4.19. Выходные статические характеристики транзистора, включенного по схеме с общим эмиттером, при больших напряжениях на коллекторе

Рис. 4.20. Статические характеристики передачи тока транзистора, включенного по схеме с общим эмиттером

чем $U_{КЭО\,проб}$, соответствующее $I_Б = 0$. В этих условиях пробой транзистора происходит при напряжениях, бо́льших $U_{КЭО\,проб}$, но не превышающих $U_{КБО\,проб}$.

Однако после того как наступил пробой, ток через эмиттерный переход существенно возрастает, что является причиной роста коэффициента передачи тока эмиттера. Теперь для продолжающегося увеличения тока коллектора, т. е. для выполнения условия $|h_{21Б}| = 1$, можно уменьшить коэффициент размножения M, уменьшив напряжение на коллекторе. Коэффициент передачи тока эмиттера при этом сохраняется равным единице, так как растет ток эмиттера. Таким образом, на статических характеристиках появляются участки с отрицательным дифференциальным сопротивлением (рис. 4.19).

Характеристики передачи тока. Семейство статических характеристик передачи тока транзистора, включенного по схеме с общим эмиттером, показано на рис. 4.20.

Общий характер этих зависимостей свидетельствует о том, что в транзисторе $I_К = I_{КЭО} + h_{21Э} I_Б$. В связи с большей зависимостью коэффициента передачи тока базы $h_{21Э}$ от режима работы транзистора по сравнению с аналогичной зависимостью коэффициента передачи постоянного тока эмиттера характеристики передачи тока в схеме с общим эмиттером отличаются большей нелинейностью.

Смещение статических характеристик передачи тока вверх при увеличении напряжения на коллекторе также более значительно по сравиеиию со смещением подобных характеристик схемы с общей базой, так как связаио в даниом случае не с умеиьшением рекомбииации дырок при уменьшеиии толщины базы, а с увеличением тока эмиттера при постоянном токе базы (см. рис. 4.18, *б*).

При $U_{КЭ} \neq 0$ характеристики передачи тока выходят не из начала координат, а из точек на оси ординат, соответствующих токам $I_{КЭО}$.

Характеристики обратиой связи. Семейство статических характеристик обратной связи транзистора, включенного по схеме с общим эмиттером, приведено иа рис. 4.21.

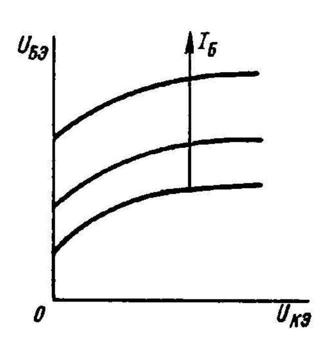


Рис. 4.21. Статические характеристики обратной связи транзистора, включенного по схеме с общим эмиттером

Общий характер этнх зависимостей можно выясиить путем аиализа рис. 4.18, *б* либо простым графическим перестроением семейства входных статических характеристик схемы с общим эмиттером.

Смещение статических характеристик обратной связи вверх при увеличении тока базы также не требует специальиого пояснения.

Больший интерес представляют статические характеристики схемы с общим эмиттером, так как в них указываются значения токов базы. Для схемы с общей базой ток базы ие измеряется, а определить его по токам эмиттера и коллектора с достаточной точностью невозможно, так как токи эмиттера и коллектора мало отличаются.

§ 4.9. РАБОТА ТРАНЗИСТОРА НА МАЛОМ ПЕРЕМЕННОМ СИГНАЛЕ

Рассмотрим подробнее факторы, влияющие на работу траизистора при условии, что на его переходы помнмо постояиных поданы малые переменные иапряжеиия. При этом в транзисторе проходят постояиные и малые перемеиные токи. Под малыми переменными иапряжеииями и токами будем понимать такие, при которых связь между ними остается линейной.

Как отмечалось в § 4.1, при работе траизистора на малом переменном сигиале его свойства определяются процессами переноса, накопления и рассасывания носителей, а также перезарядкой емкостей переходов. Чтобы рассмотреть это подробнее, представим себе, что через отдельные элементы траизистора проходит сигнал в виде иебольшого прямоугольиого импульса тока либо иебольшого тоикого пакета иосителей заряда.

Барьерная емкость эмиттера. Пусть прямоугольный импульс тока проходит через цепь эмиттера транзистора, включенного по схеме с общей базой. Исходя из принципа работы транзистора, для того чтобы получить усиление, в базу должны быть инжектированы носители заряда, для чего требуется изменение напряжения на эмиттерном переходе. Однако в связи с тем, что эмиттерный переход обладает барьерной емкостью $C_{э.бар}$, а цепь, в которую он включен, — конечным сопротивлением, напряжение на эмиттерном переходе изменится не мгновенно; значит, не

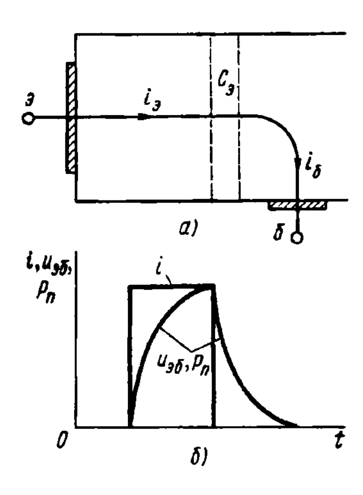


Рис. 4.22. Искажение сигнала из-за влияния барьерной емкости эмиттера:

а — путь емкостной составляющей тока; *б* — форма импульса тока и напряжения на эмиттерном переходе (граничная концентрация неосновных носителей заряда в базе около эмиттерного перехода)

Рис. 4.23. Искажение сигнала при прохождении неосновных носителей заряда через базу транзистора:

а — схематическая структура транзистора; *б* — движение пакета неосновных носителей заряда в базе; *в* — импульсы токов

мгновенно изменяется и инжекция носителей заряда (рис. 4.22). Таким образом, наличие емкости эмиттера приводит к тому, что инжекция носителей в базу как бы замедляется, а если сигналы следуют один за другим через интервалы времени, сравнимые с этим замедлением, соответствующие пакеты инжектированных носителей заряда сливаются. Это первая причина искажения высокочастотных сигналов.

Пролет носителей заряда через базу. Представим себе, что со стороны эмиттера в базу входит пакет инжектированных носителей заряда (рис. 4.23). Эти носители нарушают нейтральность базы, и в нее из невыпрямляющего контакта должно войти соответствующее число основных носителей. Это происходит довольно быстро — за время диэлектрической релаксации. Следовательно, через вывод базы проходит импульс тока, равный им-

пульсу тока эмиттера. Далее пакет носителей начинает двигаться в базе к коллектору. Если в базе носители не рекомбинируют и поле там неизменно, при таком движении отсутствует ток во всех электродах транзистора, так как инжектированный в базу заряд нейтрализован.

За время движения через базу пакет носителей заряда размывается из-за диффузии. Действительно, наличие градиентов концентрации приводит к тому, что носители заряда, находящиеся на переднем крае такого пакета, обгоняют его, а на заднем крае отстают и к коллектору подходит уже размытый пакет носителей.

Размытие пакета носителей заряда усиливается тем, что они проходят в транзисторе разные пути из-за непараллельности p-n-переходов эмиттера и коллектора. Особенно сильно это явление сказывается у краев эмиттера.

Рис. 4.24. Искажение сигнала при прохождении носителей заряда через коллекторный переход:

а — структура коллекторного перехода и прилегающих к нему областей; *б* — распределение носителей в коллекторном переходе; *в* — импульсы тока коллектора (*1* — без учета барьерной емкости коллекторного перехода; *2* — с учетом перезаряда барьерной емкости коллекторного перехода)

Когда пакет носителей доходит до коллекторного перехода, неосновные носители заряда вытягиваются из базы, ее электрическая нейтральность нарушается; следовательно, основные носители тоже выходят из базы. Если бы в базе не происходила рекомбинация, вышедший из базы заряд оказался бы равным вошедшему, т. е. среднее значение тока базы оказалось бы равным нулю. Однако через вывод базы протекает переменный ток, связанный с изменением заряда носителей в базе. Так как в базе происходит рекомбинация, выходящий из нее заряд несколько меньше входящего, что и дает некоторое ненулевое значение среднего тока базы.

Когда время пролета носителей заряда через базу $t_{прол}$ оказывается большим, размытие импульса тоже увеличивается. Тогда, если сигнал представляет собой последовательность импульсов, они перестают различаться в токе коллектора. Это вторая причина искажения высокочастотных сигналов транзистором.

Пролет носителей через область объемного заряда коллектора. Рассмотрим, как появляется ток в цепи коллектора, когда в область объемного заряда коллекторного перехода входит пакет носителей заряда (рис. 4.24). Природа проходящего при этом тока состоит в том, что вошедшие в область объемного заряда носите-

ли искажают распределение электрического поля в нем. Изменение электрического поля приводит к появлению токов смещения, которые замыкаются через цепь коллектора. Такое изменение электрического поля, а значит, и прохождение тока в цепи коллектора имеет место до тех пор, пока носители находятся в области объемного заряда (см. § 3.28). Следовательно, длительность импульса тока в цепи коллектора определяется временем пролета носителей заряда через *p-n*-переход коллектора (рис. 4.24, *в*). Если же сигналы следуют с интервалами времени, меньшими, чем время пролета через *p-n*-переход коллектора, они оказываются неразличимыми. Это третья причина искажения высокочастотных сигналов транзистором.

Барьерная емкость коллекторного перехода. Приведенные рассуждения основывались на предположении, что как только носители заряда попадают в коллекторный переход, сразу в цепи коллектора появляется ток. Однако этот ток, как отмечалось, может появиться, если изменится электрическое поле в *p-n*-переходе коллектора. Но так как коллекторный переход обладает некоторой емкостью, а цепи, в которые он включен, — сопротивлением, появление тока происходит не мгновенно — сначала должно произойти изменение заряда барьерной емкости перехода. Это также приводит к искажению сигнала (рис. 4.24, *в*).

Таким образом, работа транзистора на переменном сигнале определяется как прохождением активных токов, аналогичных по физической природе постоянным токам, так и прохождением емкостных токов, связанных с наличием барьерных емкостей и накоплением зарядов в базе. На все эти явления оказывают влияние времена пролета носителей заряда. Следовательно, для малого переменного сигнала транзистор представляет собой довольно сложный активный элемент электрической схемы, параметры которого оказываются комплексными и зависящими от частоты.

§ 4.10. МАЛОСИГНАЛЬНЫЕ ПАРАМЕТРЫ

Если переменные напряжения на переходах транзистора достаточно малы, токи в нем оказываются линейными функциями этих напряжений. Транзистор можно рассматривать как линейный четырехполюсник (рис. 4.25). При этом два внешних вывода четырехполюсника считают входными, соответствующие им ток и напряжение обозначают $\dot I_1$ и $\dot U_1$. Два других вывода являются выходными, соответствующие им ток и напряжение обозначают $\dot I_2$ и $\dot U_2$. За положительное принимают направление токов, входящих в четырехполюсник.

Чтобы показать связь между величинами $\dot U$ и $\dot U_2$, $\dot I_1$ и $\dot I_2$, составим шесть систем уравнений. Однако практически приме-

няют только три системы. В первой из них напряжения рассматривают как линейные функции токов:

$$\left.\begin{aligned}\dot U_1 &= z_{11}\dot I_1 + z_{12}\dot I_2;\\ \dot U_2 &= z_{21}\dot I_1 + z_{22}\dot I_2.\end{aligned}\right\} \qquad (4.55)$$

Коэффициенты z_{ik}, имеющие размерность сопротивления и являющиеся комплексными, можно выразить через токи и напряжения, измеренные в режиме холостого хода, следующим образом:

$$\left.\begin{aligned}z_{11} &= \frac{\dot U_1}{\dot I_1}\bigg|_{\dot I_2=0}; & z_{12} &= \frac{\dot U_1}{\dot I_2}\bigg|_{\dot I_1=0};\\ z_{21} &= \frac{\dot U_2}{\dot I_1}\bigg|_{\dot I_2=0}; & z_{22} &= \frac{\dot U_2}{\dot I_2}\bigg|_{\dot I_1=0}.\end{aligned}\right\} \qquad (4.56)$$

Здесь, как и в дальнейшем, индекс 11 (читается «один — один») означает входной параметр (характеризующий входную цепь), индекс 12 («один — два») — параметр обратной связи, индекс 21 («два-один») — параметр прямой передачи и индекс 22 («два-два») — выходной параметр.

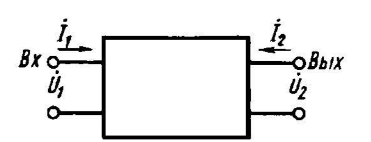


Рис. 4.25. Четырехполюсник, эквивалентный транзистору

Для получения режима холостого хода в цепь включают сопротивление, значительно большее, чем соответствующее сопротивление четырехполюсника (входное или выходное). При экспериментальном определении параметров транзистора необходимо обеспечить питание его электродов постоянным напряжением либо через очень большое активное сопротивление от достаточно высоковольтного источника питания, либо через индуктивные элементы.

Осуществление режима холостого хода в цепи эмиттера или базы (для схемы с общим эмиттером) не представляет особого труда, так как внутреннее сопротивление открытого перехода мало, а создание режима холостого хода в цепи коллектора (выходной) затруднено тем, что внутреннее сопротивление при этом очень велико (достигает нескольких мегом). Следовательно, определить экспериментально z-параметры транзистора трудно.

Если токи транзистора рассматривать как линейные функции напряжений, получаем систему уравнений:

$$\left.\begin{aligned}\dot I_1 &= y_{11}\dot U_1 + y_{12}\dot U_2;\\ \dot I_2 &= y_{21}\dot U_1 + y_{22}\dot U_2.\end{aligned}\right\} \qquad (4.57)$$

Коэффициенты y_{ik}, имеющие размерность проводимости и тоже комплексные, определяем в режиме короткого замыкания:

$$\left.\begin{aligned} y_{11} &= \left| \frac{\dot I_1}{\dot U_1} \right|_{\dot U_2=0}; \quad y_{12} = \left| \frac{\dot I_1}{\dot U_2} \right|_{\dot U_1=0}; \\ y_{21} &= \left| \frac{\dot I_2}{\dot U_1} \right|_{\dot U_2=0}; \quad y_{22} = \left| \frac{\dot I_2}{\dot U_2} \right|_{\dot U_1=0}. \end{aligned}\right\} \tag{4.58}$$

Получение режима короткого замыкания состоит в том, что исследуемую цепь шунтируют сопротивлением, значительно меньшим внутреннего сопротивления соответствующей цепи. Учитывая необходимость обеспечения питания электродов транзистора постоянным напряжением, такое шунтирование можно производить емкостью.

Режим короткого замыкания можно легко осуществить в цепи коллектора, где внутреннее сопротивление транзистора велико. Выполнение режима короткого замыкания в цепи эмиттера затруднено, особенно в области низких частот. Это ограничивает применение y-параметров для описания свойств транзистора.

Во избежание перечисленных трудностей чаще всего для описания свойств транзистора используют так называемую смешанную систему

$$\left.\begin{aligned} \dot U_1 &= h_{11}\dot I_1 + h_{12}\dot U_2; \\ \dot I_2 &= h_{21}\dot I_1 + h_{22}\dot U_2. \end{aligned}\right\} \tag{4.59}$$

Для определения h-параметров необходимы режим короткого замыкания в выходной цепи и режим холостого хода во входной. Физический смысл h-параметров довольно прост:

$h_{11} = \left.\frac{\dot U_1}{\dot I_1}\right|_{\dot U_2=0}$ — входное сопротивление при коротком замыкании выходной цепи;

$h_{12} = \left.\frac{\dot U_1}{\dot U_2}\right|_{\dot I_1=0}$ — коэффициент обратной связи по напряжению при холостом ходе во входной цепи;

$h_{21} = \left.\frac{\dot I_2}{\dot I_1}\right|_{\dot U_2=0}$ — коэффициент передачи тока при коротком замыкании выходной цепи;

$h_{22} = \left.\frac{\dot I_2}{\dot U_2}\right|_{\dot I=0}$ — выходная проводимость при холостом ходе во входной цепи.

Преимущества h-параметров состоят в удобстве их экспериментального определения. Кроме того, h-параметры измеряют в режимах, близких к режимам работы транзисторов в практических схемах. Однако для расчета электрических схем часто удобнее применять другие параметры (например, y).

Переход от одной системы параметров к другой довольно прост. С этой целью уравнения системы, от которой осуществляется переход, следует решить относительно величин, являющихся функциями в системе, к которой осуществляется переход. Полученные коэффициенты при токах или напряжениях и дадут формулы перехода. Например, если необходимо перейти от сис-

темы y-параметров к системе z-параметров, уравнения (4.57) надо решить относительно $\dot{U}_1$ и $\dot{U}_2$ (функции в системе z). Полученные выражения имеют вид:

$$\left.\begin{aligned} \dot{U}_1 &= \frac{y_{22}}{\Delta_y} \dot{I}_1 - \frac{y_{12}}{\Delta_y} \dot{I}_2; \\ \dot{U}_2 &= -\frac{y_{21}}{\Delta_y} \dot{I}_1 + \frac{y_{11}}{\Delta_y} \dot{I}_2, \end{aligned}\right\} \qquad (4.60)$$

где $\Delta_y = y_{11}y_{22} - y_{21}y_{21}$ — определитель системы y.

Сравнив систему (4.60) с (4.55), получим соотношения для перехода от одной системы параметров к другой. Найденные аналогичным путем формулы сведены в табл. 4.1.

Таблица 4.1

Формулы перехода между системами параметров

	$\|z\|$	$\|y\|$	$\|h\|$
$\|z\|$	z_{11} $\quad z_{12}$ z_{21} $\quad z_{22}$	$-\frac{y_{22}}{\Delta_y}$ $\quad -\frac{y_{12}}{\Delta_y}$ $-\frac{y_{21}}{\Delta_y}$ $\quad \frac{y_{11}}{\Delta_y}$	$\frac{\Delta_h}{h_{22}}$ $\quad \frac{h_{12}}{h_{22}}$ $-\frac{h_{21}}{h_{22}}$ $\quad \frac{1}{h_{22}}$
$\|y\|$	$\frac{z_{22}}{\Delta_z}$ $\quad -\frac{z_{12}}{\Delta_z}$ $-\frac{z_{21}}{\Delta_z}$ $\quad \frac{z_{11}}{\Delta_z}$	y_{11} $\quad y_{12}$ y_{21} $\quad y_{22}$	$\frac{1}{h_{11}}$ $\quad -\frac{h_{12}}{h_{11}}$ $\frac{h_{21}}{h_{11}}$ $\quad \frac{\Delta_h}{h_{11}}$
$\|h\|$	$\frac{\Delta_z}{z_{22}}$ $\quad \frac{z_{12}}{z_{22}}$ $-\frac{z_{21}}{z_{22}}$ $\quad \frac{1}{z_{22}}$	$\frac{1}{y_{11}}$ $\quad -\frac{y_{12}}{y_{11}}$ $\frac{y_{21}}{y_{11}}$ $\quad \frac{\Delta_y}{y_{11}}$	h_{11} $\quad h_{12}$ h_{21} $\quad h_{22}$

Значения параметров транзистора, представленного в виде четырехполюсника, зависят от схемы его включения. Однако, если эти параметры известны для какой-либо одной схемы, сравнительно легко произвести пересчет для другой. Для этого надо заменить напряжения и токи (имея в виде правило знаков), учитывая, что в транзисторе

$$\dot{I}_э + \dot{I}_б + \dot{I}_к = 0; \qquad (4.61)$$

$$\dot{U}_{бэ} + \dot{U}_{кб} + \dot{U}_{эк} = 0. \qquad (4.62)$$

Сделав необходимые подстановки и преобразовав уравнения, получаем формулы перехода как коэффициенты в уравнениях. Например, если известны y-параметры транзистора для схемы с

общей базой, а требуется найти их для схемы с общим эмиттером, в уравнения

$$\dot{I}_э = y_{11б}\dot{U}_{эб} + y_{12б}\dot{U}_{кб}; \qquad (4.63)$$

$$\dot{I}_к = y_{21б}\dot{U}_{эб} + y_{22б}\dot{U}_{кб} \qquad (4.64)$$

подставляем токи и напряжения с помощью выражений (4.61) и (4.62). После преобразований получим

$$\dot{I}_б = (y_{11б} + y_{12б} + y_{21б} + y_{22б})\,\dot{U}_{бэ} - (y_{12б} + y_{22б})\dot{U}_{кэ}, \qquad (4.65)$$

$$\dot{I}_к = -(y_{21б} + y_{22б})\dot{U}_{бэ} + y_{22б}\dot{U}_{кэ}, \qquad (4.66)$$

где $\dot{U}_{бэ} = -\dot{U}_{эб}$.

Следовательно,

$$\left.\begin{aligned} y_{11э} &= y_{11б} + y_{12б} + y_{21б} + y_{22б};\\ y_{12э} &= -(y_{12б} + y_{22б});\\ y_{21э} &= -(y_{21б} + y_{22б});\\ y_{22э} &= y_{22б}. \end{aligned}\right\} \qquad (4.67)$$

Аналогично выводят соотношения и для других параметров.

§ 4.11. ЭКВИВАЛЕНТНЫЕ СХЕМЫ

Рассмотрение транзистора как активного линейного четырехполюсника (см. § 4.10) удобно для расчета электрических схем. Однако оно имеет и ряд недостатков, которые связаны прежде всего с тем, что параметры четырехполюсника вводят в известной степени формально и каждый из них может отражать влияние сразу нескольких физических процессов. Поэтому получаются сложные зависимости параметров четырехполюсника от режима работы транзистора (постоянных напряжений и токов), от частоты и температуры. Чтобы упростить эти зависимости, свойства транзистора при малом переменном сигнале описывают с помощью эквивалентных схем. Под эквивалентной понимают электрическую схему, составленную из линейных элементов электрических цепей (сопротивлений, емкостей, индуктивностей, генераторов тока или напряжения), которая по своим свойствам при данном сигнале (например, при малом переменном) не отличается от реального объекта (транзистора). Графическое изображение эквивалентных схем позволяет более экономно зафиксировать основные соотношения. При расчетах с помощью эквивалентных схем сначала определяют токи и напряжения в самой схеме и затем переходят к каким-то другим параметрам, например параметрам четырехполюсника.

Никакая эквивалентная схема из конечного числа элементов не может быть полностью эквивалентной реальному транзистору, т. е. все эквивалентные схемы оказываются приближенными. Чем проще эквивалентная схема, чем меньше она содержит эле-

ментов, тем легче ею пользоваться, но обычно тем менее точно она отражает свойства реального транзистора.

По способу построения различают формальные и фнзические эквивалентиые схемы.

Формальные эквивалентные схемы строят на осиове описання транзистора с помощью уравиений четырехполюсника (рнс. 4.26). Как видно из рисуика, каждая схема содержит по четыре элемента: два сопротивлеиия (комплексных) и два геиератора тока или напряжения. Такие эквнвалеитиые схемы не имеют никаких преимуществ по сравнеиию с описанием траизистора как четы-

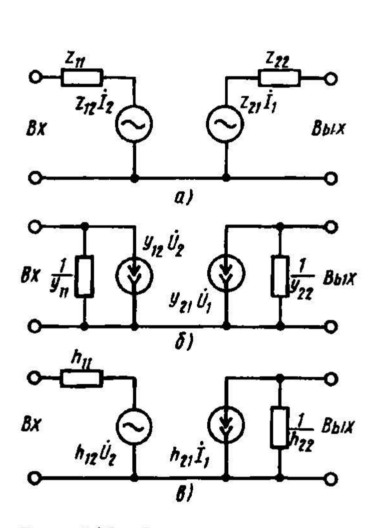


Рис. 4.26. Формальные эквивалентные схемы транзистора:

а — для системы *z*-параметров; *б* — для системы *y*-параметров; *в* — для системы *h*-параметров

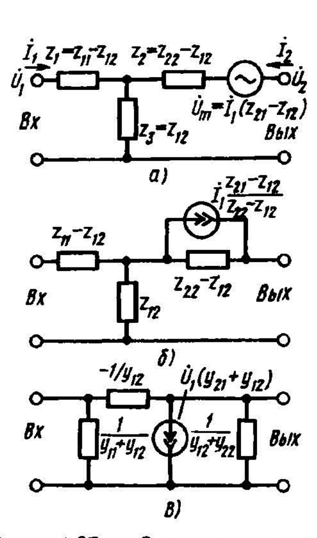


Рис. 4.27. Одиогенераториые формальные эквивалентные схемы:

а — Т-образиая эквивалентная схема с генератором ЭДС; *б* — Т-образиая эквивалеитиая схема с генератором тока; *в* — П-образиая эквивалентная схема с генератором тока

рехполюсника путем задания его параметров (нлн соответствующих уравнений).

Формальные эквивалеитные схемы можно представнть в разной форме: например, построить так, чтобы в схеме был только один активный элемент (рис. 4.27). Обычно этн схемы по способу начертання подразделяют на Т- и П-образиые. Чтобы установить связь элементов одногенераториых формальиых схем с параметрами четырехполюсника, можно записать выражения для токов н иапряженнй в ннх, а затем сопоставить со зиачеииями токов и иапряжеиий четырехполюсиика. Например, для Т-образиой эквивалентиой схемы с генератором ЭДС (рис. 4.27, *а*) в режиме холостого хода иа выходе

$$\left.\begin{aligned}\dot U_1 &= \dot I_1(z_1 + z_3),\\ \dot U_2 &= \dot U_m + \dot I_1 z_3,\end{aligned}\right\} \tag{4.68}$$

в режиме холостого хода на входе

$$\left.\begin{aligned}\dot U_1 &= \dot I_2 z_3,\\ \dot U_2 &= \dot I_2(z_2 + z_3).\end{aligned}\right\} \tag{4.69}$$

Отсюда

$$\left.\begin{aligned}&\frac{\dot U_2}{\dot I_1}\bigg|_{\dot I_2=0} = z_3 + \frac{\dot U_m}{\dot I_1} = z_{21};\ \frac{\dot U_1}{\dot I_1}\bigg|_{\dot I_2=0} = z_1 + z_2 = z_{11};\\ &\frac{\dot U_1}{\dot I_2}\bigg|_{\dot I_1=0} = z_3 = z_{12};\ \frac{\dot U}{\dot I_2}\bigg|_{\dot I_1=0} = z_2 + z_3 = z_{22}.\end{aligned}\right\} \tag{4.70}$$

После преобразования получим значения, указанные на рисунке. Аналогичио получают параметры и других одиогенераториых схем.

Физические эквивалентные схемы составляют на основании физических соображений для определенных типов конструкций транзистора, для определенного частотного диапазоиа, ориентируясь на определенную схему включения транзистора (с общим эмиттером, общей базой, общим коллектором). Каждый вывод физической эквивалентной схемы соответствует электроду транзистора. Заметим, что в формальных эквивалентных схемах различают только входные и выходные зажимы независимо от того, какими электродами траизистора оии являются.

Физическую эквивалеитную схему строят, выделяя мысленно иекоторые части в транзисторе и рассматривая отдельно процессы в этих частях.

За основу построения обычио берут формальную эквивалентиую схему идеализнрованного транзистора, называемого *одномерной теоретической моделью*. При изучеиии одномерной теоретнческой модели траизистора считают, что носители заряда в ием движутся по параллельным путям, а поверхностная рекомбииация только изменяет время жизнн носителей. Кроме того, в одномериой теоретической модели не учитывают влияние объемиых сопротивлений и токов, проходящих через барьерные емкости переходов. При таких предположениях получается, что параметры формальиой эквивалентиой схемы довольно просто выразить через конструктнвные параметры идеализированного транзистора (толщиину базы), режим его работы, свойства материала.

К одномерной теоретической модели транзистора добавляют элемеиты, учитывающие другие процессы, например падения напряжений иа объемных сопротивлениях, токи через емкости и т. д.

§ 4.12. ЭКВИВАЛЕНТНАЯ СХЕМА ОДНОМЕРНОЙ ТЕОРЕТИЧЕСКОЙ МОДЕЛИ

Как отмечалось, параметры формальной эквивалентной схемы легко можно выразить через параметры четырехполюсника, которые в свою очередь могут быть найдены по значениям токов и напряжений в транзисторе. Поэтому для построения формальной эквивалентной схемы одномерной теоретической модели транзистора нужно знать переменные составляющие токов и напряжений. Строгий расчет этих составляющих производится примерно так же, как и расчет постоянных составляющих (см. § 4.4). Однако он сложнее, так как в данном случае менее обоснованно предположение о постоянстве плотности тока во всех сечениях базы, особенно сильно это проявляется в области высоких частот.

Для дальнейших расчетов воспользуемся упрощенным приемом. Учитывая, что все параметры формальной эквивалентной схемы комплексные, изобразим их как результат параллельного соединения сопротивлений и емкостей. Полученная эквивалентная схема показана на рис. 4.28. Эта эквивалентная схема должна быть справедлива в области низких частот. Тогда активные сопротивления эквивалентной схемы можно получить как отношения приращений напряжений в цепях транзистора к вызвавшим их приращениям токов (такие приращения играют на низких частотах роль переменных сигналов).

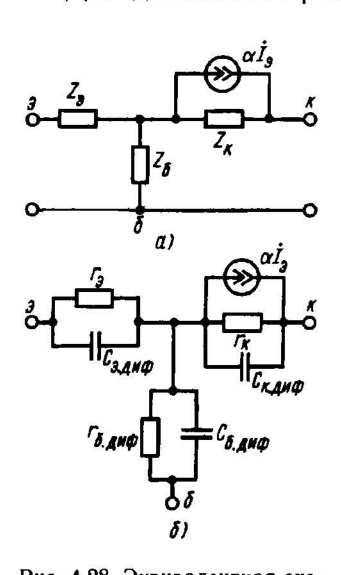


Рис. 4.28. Эквивалентная схема одномерной теоретической модели транзистора:
а — с комплексными сопротивлениями; *б* — с активными сопротивлениями и диффузионными емкостями

Согласно эквивалентной схеме, *сопротивление эмиттера*

$$r_э = \left.\frac{du_{эб}}{di_э}\right|_{i_к=\text{const}} - \left.\frac{du_{эб}}{di_к}\right|_{i_э=\text{const}}; \quad (4.71)$$

сопротивление коллектора

$$r_к \approx \left.\frac{du_{кб}}{di_к}\right|_{i_э=\text{const}}; \quad (4.72)$$

диффузионное сопротивление базы

$$r_{б.диф} = \left.\frac{du_{эб}}{di_к}\right|_{i_э=\text{const}}; \quad (4.73)$$

коэффициент передачи тока эмиттера одномерной теоретической модели транзистора

$$\alpha = \left.\frac{di_к}{di_э}\right|_{u_{кб}=\text{const}}. \quad (4.74)$$

Емкости в эквивалентной схеме одномерной теоретической модели отражают только накопление носителей заряда в базе. Диффузионные емкости транзистора

$$\frac{1}{C_{\text{э диф}}} = \frac{du_{\text{эб}}}{dQ}\bigg|_{i_{\text{к}}=\text{const}} - \frac{du_{\text{эб}}}{dQ}\bigg|_{i_{\text{э}}=\text{const}}; \tag{4.75}$$

$$C_{\text{б диф}} = \frac{dQ}{du_{\text{эб}}}\bigg|_{i_{\text{э}}=\text{const}}; \tag{4.76}$$

$$\frac{1}{C_{\text{к. диф}}} = \left(\frac{du_{\text{кб}}}{dQ} - \frac{du_{\text{эб}}}{dQ}\right)\bigg|_{i_{\text{э}}=\text{const}} \tag{4.77}$$

Справедливость соотношений (4.71) — (4.77) сохраняется вплоть до частот, когда период сигнала иачнет приближаться ко времени пролета носителей заряда через базу.

Рассмотрим, какие процессы в базе одномерной теоретической модели транзистора определяют значения параметров ее эквивалентной схемы.

Генератор тока в цепи коллектора учитывает, как указывалось, активное свойство транзистора. Значение тока этого генератора пропорционально значению тока эмиттера и зависит от частоты. При этом изменяются как амплитуда, так и фаза тока.

Частотные зависимости α подробно рассмотрены в § 4.14, здесь ограничимся изучением его низкочастотного значения α. Воспользовавшись выражениями для токов транзистора (см. § 4.4), с помощью формулы (4.74) можно найти значение α. Чтобы знать, какими процессами определяется значение α, его представляют в виде произведения четырех сомножителей:

$$\alpha = \gamma\alpha_{\text{п}}\alpha^{*}M, \tag{4.78}$$

где $\gamma = \frac{di_{\text{эр}}}{di_{\text{э}}}\Big|_{u_{\text{кб}}=\text{const}}$ — *эффективность эмиттера,* показывающая, какова доля инжектированных в базу носителей заряда в общем токе эмиттера (значение γ обычно немного меньше единицы); $\alpha_{\text{п}} = \frac{di_{\text{кр}}}{di_{\text{эр}}}\Big|_{u_{\text{кб}}=\text{const}}$ — *коэффициент переноса,* показывающий, какая часть инжектированных в базу носителей заряда доходит до коллекторного перехода (значение $\alpha_{\text{п}}$ обычно немного меньше единицы; $\alpha_{*} = \frac{di_{\text{к}}}{di_{\text{кр}}}\Big|_{\substack{u_{\text{кб}}=\text{const}\\ M=1}}$ — *эффективность коллектора,* показывающая, во сколько раз ток коллектора возрастает из-за наличия в обратном токе коллекторного перехода составляющей, зависящей от тока эмиттера (значение α^{*} обычно несколько больше единицы или равно единице); M — коэффициент лавинного размножения (см. § 3.11).

Эффективность коллектора α^{*} учитывает явление, связанное с изменением экстракции неосновных носителей из коллекторной области: при увеличении тока эмиттера и, следовательно, тока коллектора увеличивается падение напряжения на объемном

сопротивлении коллекторной области, которое способствует движению неосновных носителей коллекторной области к коллекторному переходу. В результате при увеличении тока эмиттера ток коллектора возрастает не только из-за прохождения через коллекторный переход носителей, инжектированных эмиттером, но и из-за экстракции неосновных носителей заряда из более удаленных частей коллекторной области.

Практически $\alpha^* > 1$ только в германиевых транзисторах с высокоомной коллекторной областью. В кремниевых транзисторах $\alpha^* \approx 1$, так как обратный ток через кремниевый p-n-переход обусловлен в основном тепловой генерацией носителей в самом p-n-переходе, а не экстракцией неосновных носителей заряда.

Общие соотношения, выведенные ранее, позволяют получить все сомножители (4.78) — внутренние параметры транзистора. Для бездрейфового транзистора

$$\gamma \approx 1 - \frac{n_{p0э} D_n w}{p_{n0б} D_p L_n}; \qquad (4.79)$$

$$\alpha_п \approx 1 - \frac{w^2}{2L_p^2}; \qquad (4.80)$$

$$\alpha^* \approx 1. \qquad (4.81)$$

С параметром α *связан коэффициент передачи тока базы одномерной теоретической модели транзистора*

$$\beta = \left.\frac{di_к}{di_б}\right|_{u_{кэ}=\text{const}} \approx \frac{\alpha}{1-\alpha}. \qquad (4.82)$$

Равенство (4.82) является приблизительным, так как определение коэффициентов передачи тока эмиттера α и тока базы β производится при несколько отличных условиях: $u_{кб} = \text{const}$ и $u_{кэ} = \text{const}$.

Параметры эквивалентной схемы одномерной теоретической модели транзистора α и β легко выразить через параметры эквивалентного четырехполюсника, т. е. через параметры реального транзистора:

$$\alpha \approx -h_{21б}; \qquad (4.83)$$

$$\beta \approx h_{21э}. \qquad (4.84)$$

Эти равенства также являются приблизительными, так как для одномерной теоретической модели транзистора принят ряд допущений (см. § 4.11). Знаки у коэффициентов передачи тока эмиттера одномерной теоретической модели транзистора и реального транзистора (4.83) различны в связи с отличием направлений токов в реальном транзисторе, включенном по схеме с общей базой, и условно принятыми направлениями тех же токов в эквивалентном четырехполюснике.

Сопротивление эмиттера. Если $\gamma = 1$, то сопротивление эмиттера

$$r_э = \frac{kT}{qI_э}\left(1 - \frac{t_{прол}}{\int\limits_0^w \frac{dx}{N} \int\limits_0^w \frac{N}{D_p} dx}\right). \quad (4.85)$$

Для бездрейфового транзистора, воспользовавшись соотношениями § 4.4, получим

$$r_э = \frac{kT}{2qI_э}. \quad (4.86)$$

Для дрейфового транзистора с постоянным полем в области базы

$$r_э = \frac{kT}{qI_э}\left[1 - \frac{N(w)}{N(0)} \frac{\ln\frac{N(0)}{N(w)} - \left(1 - \frac{N(w)}{N(0)}\right)}{\left(1 - \frac{N(w)}{N(0)}\right)^2}\right] \approx \frac{kT}{qI_э}. \quad (4.87)$$

Таким образом, сопротивление эмиттера у бездрейфового транзистора меньше, чем у дрейфового, примерно вдвое. Это обусловлено тем, что сопротивление эмиттера в бездрейфовых транзисторах отражает не только изменение тока эмиттера при изменении напряжения, но и учитывает влияние модуляции толщины базы. В дрейфовых транзисторах ток эмиттера в значительной степени определяется электрическим полем; следовательно, роль модуляции толщины базы в данном случае меньше.

Диффузионная емкость эмиттера. Аналогично находят диффузионную емкость эмиттера:

$$C_{э.диф} \approx \frac{qI_э}{kT} t_{прол}. \quad (4.88)$$

Диффузионное сопротивление базы. Расчет дает

$$r_{б.диф} = \frac{kT}{qI_э} \frac{\tau_p}{\int\limits_0^w \frac{dx}{N} \int\limits_0^w \frac{N}{D_p} dx}. \quad (4.89)$$

Для бездрейфового транзистора

$$r_{б.диф} = \frac{kT}{qI_э} \frac{L^2}{w^2} \approx \beta r_э, \quad (4.90)$$

для дрейфового

$$r_{б.диф} = \frac{kT}{qI_э} \frac{\ln^2 \frac{N(0)}{N(w)} L^2}{\frac{N(0)}{N(w)} w^2 \left(1 - \frac{N(w)}{N(0)}\right)^2} \approx \frac{kT}{qI_э} \frac{L^2}{w^2} \frac{N(w)}{N(0)} \ln^2 \frac{N(0)}{N(w)}. \quad (4.91)$$

Полученные результаты показывают, что для дрейфовых транзисторов диффузионное сопротивление базы меньше, чем для

бездрейфовых. Это также связано с тем, что влияние модуляции толщины базы на ток эмиттера в дрейфовых траизисторах незначнтельно.

Диффузионная емкость базы. По аналогии с определением диффузионной емкости эмиттера,

$$C_{б.диф} = \frac{qI_э}{kT} \int_0^w \frac{dx}{N} \int_0^w \frac{N}{D_p} dx. \tag{4.92}$$

Для бездрейфового транзистора

$$C_{б.диф} = \frac{qI_э}{kT} \frac{w^2}{D_p}, \tag{4.93}$$

для дрейфового

$$C_{б.диф} \approx \frac{qI_э}{kT} \frac{w^2}{D_p} \frac{N(0)}{N(w)} \frac{1}{\ln^2 \frac{N(0)}{N(w)}}. \tag{4.94}$$

Диффузионная емкость базы дрейфового транзистора получилась довольно большой, что отражает малую обратную связь между цепями коллектора и эмиттера из-за малого влияния модуляции толщины базы.

Сопротивление коллектора. Исходя из выражений для токов, получим значения сопротивлений коллектора бездрейфового транзистора

$$r_к = \frac{L^2}{I_э w \frac{dw}{du_{кб}}} \tag{4.95}$$

и дрейфового транзистора

$$r_к = \frac{L^2 \ln \frac{N(0)}{N(w)}}{I_э w \frac{dw}{du_{кб}}}. \tag{4.96}$$

Эти соотношения можно преобразовать, воспользовавшись выражением для толщины коллекторного перехода. В бездрейфовых транзисторах обычно используются резкие p-n-переходы. В этом случае

$$\frac{dw}{du_{кб}} = -\frac{d\delta_к}{du_{кб}} = \frac{\delta_к}{2U_{КБ}} \tag{4.97}$$

и

$$r_к = \frac{U_{КБ} \beta w}{I_э \delta_к}. \tag{4.98}$$

Если принять, что коллекторный переход дрейфового транзистора плавный с линейным распределением концентрации примесей, то

$$\frac{dw}{du_{кб}}=\frac{\delta_к}{3U_{КБ}} \tag{4.99}$$

и

$$r_к=\frac{3U_{КБ}\beta w}{I_Э\delta_к}. \tag{4.100}$$

Из полученных выражений следует, что сопротивление коллектора дрейфового транзистора больше, чем бездрейфового, как из-за меньшего влияния модуляции толщины базы на ток коллектора, так и из-за более слабой зависимости толщины области объемного заряда от напряжения для плавных переходов.

Диффузионная емкость коллектора. Этот параметр эквивалентной схемы определяется аналогично предыдущим:

$$C_{к.диф}=I_Э\frac{N(w)}{D_p}\int_0^w\frac{dx}{N}\frac{dw}{du_{кб}}. \tag{4.101}$$

Для бездрейфового транзистора

$$C_{к.диф}=\frac{I_Э w\delta_к}{2U_{КБ}D_p}, \tag{4.102}$$

для дрейфового

$$C_{к.диф}=\frac{I_Э\delta_к w}{3U_{КБ}D_p}\frac{1}{\ln\dfrac{N(0)}{N(w)}}. \tag{4.103}$$

§ 4.13. БАРЬЕРНЫЕ ЕМКОСТИ ПЕРЕХОДОВ И СОПРОТИВЛЕНИЕ БАЗЫ

Определение барьерных емкостей переходов транзистора в принципе ничем не отличается от расчетов, приведенных в § 2.8. При расчете барьерной емкости коллектора часто можно пренебречь контактной разностью потенциалов, учитывая, что постоянное напряжение на коллекторе довольно велико.

Расчет барьерной емкости эмиттера имеет некоторые особенности. Прежде всего обязательно надо учитывать контактную разность потенциалов, так как напряжение на эмиттере прямое. Кроме того, в цепи эмиттера часто задается ток (см. § 4.8), а не напряжение. Это связано с малым входным сопротивлением транзистора. Поэтому напряжение в формуле для барьерной емкости лучше тоже выразить через ток эмиттера.

Например, для бездрейфового транзистора с резким эмиттерным переходом

$$C_{э.бар}=S_э\sqrt{\frac{\varepsilon\varepsilon_0 qN}{2(\varphi_{кон}-U_{ЭБ})}}. \tag{4.104}$$

Контактная разность потенциалов из (2.2)

$$\varphi_{\text{кон}} = \frac{kT}{q} \ln \frac{n_{n0} p_{p0}}{n_i^2}. \qquad (4.105)$$

Напряжение на эмиттере из (4.16)

$$U_{\text{ЭБ}} = \frac{kT}{q} \ln \frac{I_{\text{Э}} w}{q p_{n0} D_p S_{\text{э}}}. \qquad (4.106)$$

Подставив эти выражения в (4.104), получим

$$C_{\text{э.бар}} = S_{\text{э}} \sqrt{\frac{\varepsilon\varepsilon_0 q N}{2 \frac{kT}{q} \ln \frac{q N_{\text{э}} D_p S_{\text{э}}}{I_{\text{Э}} w}}}. \qquad (4.107)$$

База транзистора обладает, как отмечалось, сопротивлением; следовательно, проходящие через нее токи могут создавать на

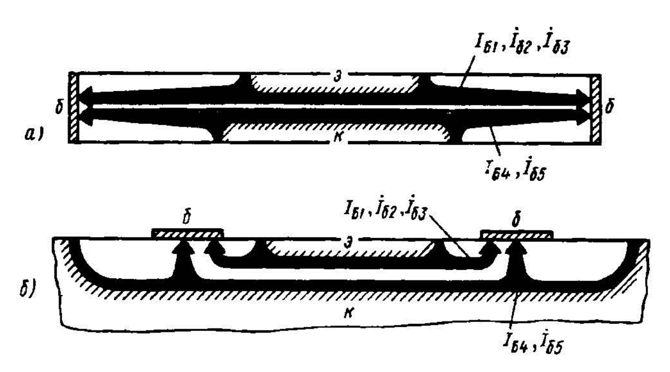


Рис. 4.29. Пути токов в базе транзисторов (схематично):
а — сплавного; *б* — планарного

этом сопротивлении падения напряжения, которые прикладываются к переходам транзистора, создавая обратные связи. Кроме того, и само сопротивление базы определяет значение постоянных времени заряда барьерных емкостей переходов. Поэтому для определения роли сопротивления базы и его места в эквивалентной схеме транзистора рассмотрим распределение токов в базе. Характер такого распределения в значительной степени зависит от конструкции транзистора (рис. 4.29).

Токи в области базы транзистора можно подразделить на следующие:

1) постоянная составляющая тока базы, связанная с рекомбинацией, с инжекцией носителей заряда из базы в эмиттер и частично с лавинным умножением тех носителей, которые идут от эмиттера в коллектор,

$$I_{\text{Б1}} \approx I_{\text{Э}} / h_{21\text{э}}; \qquad (4.108)$$

путь этой составляющей тока идет из активной части базы к выводу базы;

2) переменная составляющая тока базы, связанная с теми же процессами, что и постоянная составляющая, а также с накоплением заряда в базе,

$$\dot{I}_{б2} \approx \dot{I}_э / h_{21э}; \qquad (4.109)$$

путь этой составляющей тока тот же, что и первой;

3) переменная составляющая тока базы, связанная с зарядом барьерной емкости эмиттера,

$$\dot{I}_{б3} = j\omega C_{э.бар} \dot{U}_{эб}; \qquad (4.110)$$

эта составляющая тока проходит практически тот же путь, что и первые две;

4) обратный ток коллектора

$$I_{Б4} = I_{КБО}; \qquad (4.111)$$

эта составляющая тока базы идет от коллекторного перехода к выводу базы;

5) ток заряда барьерной емкости коллектора

$$\dot{I}_{б5} = j\omega C_{к.бар} \dot{U}_{кб}; \qquad (4.112)$$

этот ток проходит путь примерно такой же, как и обратный ток коллектора.

Влияние этих токов на работу транзистора различно. Постоянные составляющие приводят к смещению рабочих точек и перераспределению плотности тока по площади эмиттера. Кроме того, следует учесть, что сопротивление, через которое проходят постоянные токи, модулировано — под действием переменного напряжения коллектора изменяется ширина области объемного заряда коллекторного перехода, что приводит к изменению толщины базы, т. е. сечения, через которое проходят постоянные токи. Прохождение постоянных токов через переменное сопротивление приводит к проявлению переменного напряжения, которое прикладывается к p-n-переходам транзистора, приводя к появлению обратной связи.

Переменные составляющие токов, проходя через сопротивление базы, вызывают появление переменных напряжений, которые также приводят к появлению обратных связей в транзисторе. Такие обратные связи образуются по-разному в зависимости от конструкции транзистора. Так, в конструкции, представленной на рис. 4.29, *а*, к эмиттерному переходу прикладывается падение напряжения, вызванное всем током заряда барьерной емкости коллектора. При другой конструкции (рис. 4.29, *б*) только ток, идущий от части коллектора, примыкающей к активной и пассивной частям базы, создает падение напряжения, которое влияет на разность потенциалов между эмиттером и базой. Ток, идущий через периферическую часть базы, такой обратной связи не создает.

Помимо создания обратных связей в транзисторе переменные составляющие тока обеспечивают изменение зарядов барьерных

емкостей, так что соответствующие сопротивлеиия определяют постояииые времеии.

Особеиности расчета сопротивлеиий базы:

1) так как оии представляют собой распределеииые сопротивлеиия каких-то объемиых областей, необходимо учитывать растекаиие токов;

2) в связи с распределеииым характером сопротивления базы следует учитывать какой-то средиий эффект, производимый таким сопротивлеиием.

Кроме того, в траизисторах с иеравиомериым распределением примесей в базе иеобходимо учитывать еще, что удельиое сопротивлеиие в разных точках различно. Одиако база траизистора обычно представляет собой очеиь тонкий слой и ток базы проходит практически только вдоль этого слоя. Тогда вместо обычиого удельиого сопротивлеиия можио вводить сопротивление квадратиого участка слоя ϱ_s

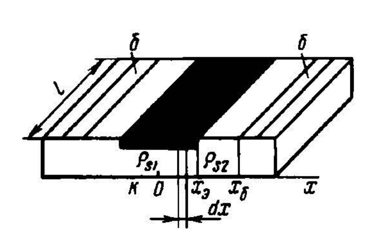


Рис. 4.30. Структура транзистора, принятая для расчета сопротивления базы

$$\frac{1}{\varrho_s}=q\int_0^w \mu N dx, \qquad (4.113)$$

аналогичное по смыслу удельному поверхностиому сопротивлеиию.

Произведем расчет коэффициента обратиой связи между цепями коллектора и эмиттера транзистора, структура которого показаиа иа рис. 4.30. Плотность емкостиого тока коллектора

$$\dot J_к = j\omega \frac{C_{к.бар}}{S_к} \dot U_{кб}. \qquad (4.114)$$

Ток с такой плотиостью вводится в базу. Пусть иа участке от 0 до $x_э$ сопротивлеиие квадрата слоя базы ϱ_{s1}. Тогда падение напряжеиия на участке dx

$$d\dot U = -\dot I \frac{\varrho_{s1}}{l} dx. \qquad (4.115)$$

Будем считать, что возиикающие в базе иапряжения невелики, так что отводимый из базы через эмиттериый переход ток тоже иевелик. Тогда ток, проходящий через сечеиие dx, составляет $\dot I = xl\dot J_к$. Проиитегрировав (4.115), на участке от $x_э$ до x получим

$$\dot U - \dot U_{x_э} = \dot J_к \frac{\varrho_{s1}}{2}(x_э^2 - x^2). \qquad (4.116)$$

Аиалогичио иаходим распределеиие напряжеиия иа участке от $x_э$ до $x_б$:

$$\dot U = \dot J_к \frac{\varrho_{s2}}{2}(x_б^2 - x^2); \qquad (4.117)$$

откуда

$$\dot U_{x_э} = \dot J_к \frac{\varrho_{s2}}{2}(x_б^2 - x_э^2). \qquad (4.118)$$

Тогда на участке от 0 до $x_э$

$$\dot U = \frac{\dot J_к}{2}\left[\varrho_{s1}(x_э^2 - x^2) + \varrho_{s2}(x_б^2 - x_э^2)\right]. \qquad (4.119)$$

Таким образом, напряжение в базе распределено неравномерно. Чтобы определить усредненное значение напряжения, создающего обратную связь, учтем, что при этом ток эмиттера должен быть равен нулю, т. е.

$$\int_0^{x_э}(\dot U_{эб} - \dot U)dx = 0. \qquad (4.120)$$

Тогда

$$\dot U_{эб} = \varrho_{s1}\frac{\dot J_к}{2x_э}\int_0^{x_э}(x_э^2 - x^2)dx + \varrho_{s2}\frac{\dot J_к}{2}(x_б^2 - x_э^2). \qquad (4.121)$$

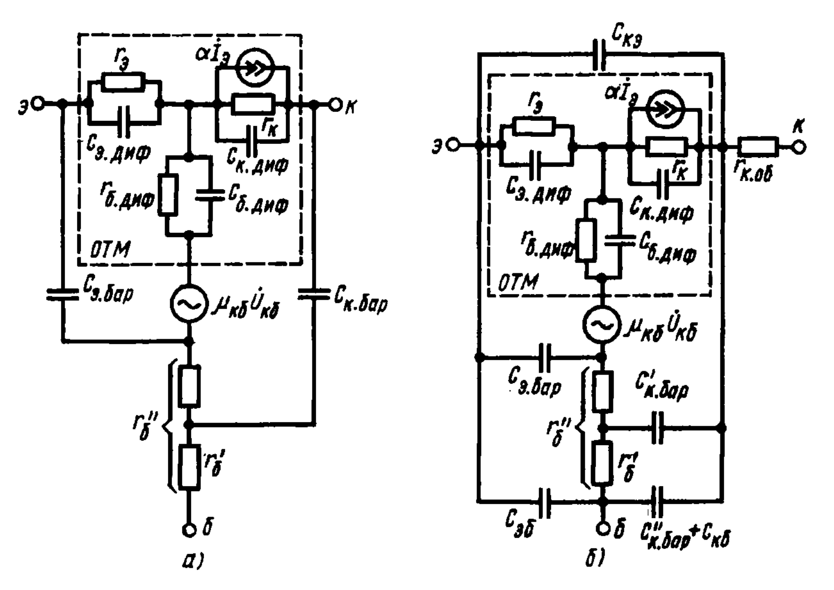


Рис. 4.31. Полные эквивалентные схемы транзисторов:
а — сплавного (рис. 4.29, *а*); *б* — планарного (рис. 4.29, *б*)

Окончательно, подставив значение $\dot J_к$, получим

$$\frac{\dot U_{эб}}{\dot U_{кб}} = j\omega C_{к.бар}\frac{1}{S_к}\left(\varrho_{s1}\frac{x_э^2}{3} + \right.$$

$$\left. + \varrho_{s2}\frac{x_б^2 - x_э^2}{2}\right) = j\omega r_б' C_{к.бар}, \qquad (4.122)$$

откуда эффективное сопротивление, определяющее обратную связь,

$$r'_б = \frac{1}{S_к}\left(Q_{s1}\frac{x_э^2}{3} + Q_{s2}\frac{x_б^2 - x_э^2}{2}\right). \qquad (4.123)$$

Аналогично определяют сопротивление базы и при других связях в транзисторе. Метод расчета нужно выбирать в зависимости от исследуемого процесса.

Рассмотрим построение полных эквивалентных схем транзисторов (рис. 4.31).

Для бездрейфового транзистора со структурой, показанной на рис. 4.29, *а*, схема имеет вид, представленный на рис. 4.31, *а*. Элементы, соответствующие одномерной теоретической модели,

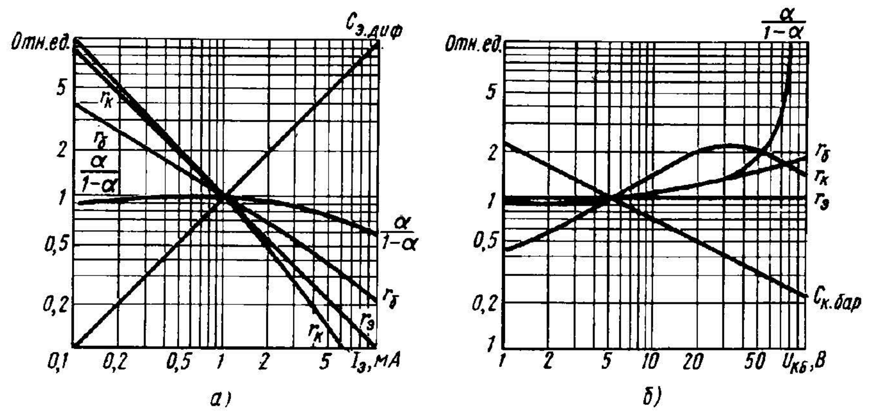


Рис. 4.32. Зависимости параметров эквивалентной схемы транзистора от постоянного тока эмиттера (*а*) и от постоянного напряжения на коллекторе (*б*)

не отличаются от представленных на рис. 4.28. В цепь базы помимо сопротивлений включен генератор ЭДС, учитывающий напряжение, которое возникает в базе при прохождении постоянного тока через переменное сопротивление.

Для дрейфового транзистора со структурой, показанной на рис. 4.29, *б*, эквивалентная схема несколько другая. Барьерная емкость коллектора в данном случае перезаряжается через разные сопротивления. Часть этой емкости $C'_{к.бар}$ создает обратную связь, а часть емкости коллектора $C''_{к.бар}$, соответствующая периферической базе, не дает обратной связи. Кроме того, из-за высоких рабочих частот, на которых работают дрейфовые транзисторы, в эквивалентной схеме целесообразно учитывать емкости между внешними выводами $C_{кэ}$, $C_{эб}$, $C_{кб}$, а также объемное сопротивление коллектора.

Режим работы транзистора (напряжение на коллекторе, ток эмиттера) существенно влияет на параметры эквивалентной

схемы. Эти зависимости объясняются соотношениями, приведенными в § 4.6, 4.12 и 4.13. На рис. 4.32 даны зависимости параметров маломощного бездрейфового германиевого транзистора от тока эмиттера (рис. 4.32, *а*) и от напряжения коллектора (рис. 4.32, *б*). Для наглядности все зависимости даются как относительные — значения параметров при $I_Э = 1$ мА и $U_{КБ} = 5$ В приняты за единицу.

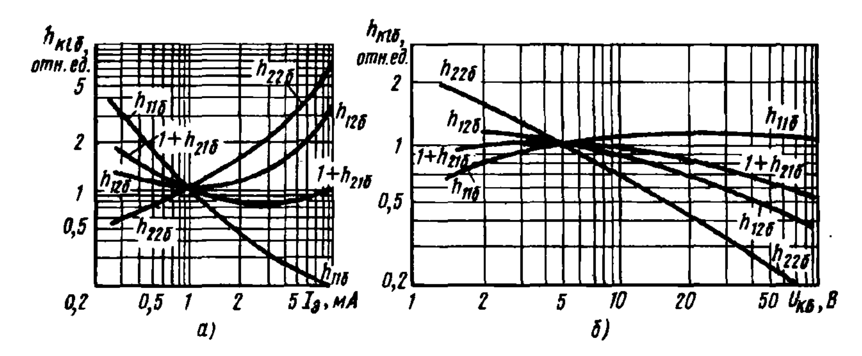


Рис. 4.33. Зависимости *h*-параметров транзистора от постоянного тока эмиттера (*а*) и от постоянного напряжения на коллекторе (*б*)

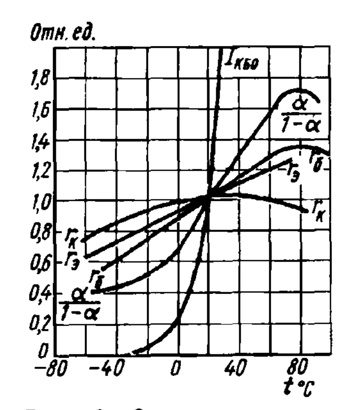


Рис. 4.34. Зависимости параметров физической эквивалентной схемы транзистора от температуры

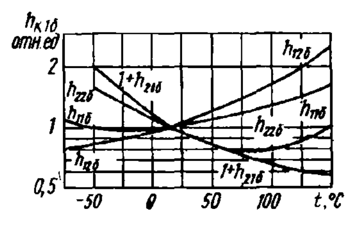


Рис. 4.35. Зависимости *h*-параметров транзистора от температуры

На рис. 4.33 показаны нормализованные зависимости *h*-параметров четырехполюсника, эквивалентного маломощному бездрейфовому германиевому транзистору, от режима работы.

У транзисторов, как и у других полупроводниковых приборов, изменяются свойства при изменении температуры, что в ряде случаев затрудняет их применение.

На рис. 4.34 приведены зависимости параметров эквивалентной схемы маломощного бездрейфового германиевого транзистора от температуры (относительные изменения значений).

На рис. 4.35 показана температурная зависимость параметров четырехполюсника.

§ 4.14. ЧАСТОТНЫЕ ХАРАКТЕРИСТИКИ

Одним из основных факторов, определяющих пригодность транзистора для использования в той или иной электрической схеме, является зависимость его параметров от частоты. Особенно большое значение имеют зависимости усилительных свойств транзистора от частоты.

Рассмотрим частотную зависимость коэффициента передачи тока эмиттера $h_{21б}$. В § 4.9 отмечалось, что на коэффициент передачи тока влияют емкость цепи эмиттера, время пролета носителей заряда через базу, время пролета носителей через область объемного заряда коллектора и постоянная времени цепи коллектора.

Постоянная времени цепи эмиттера. Как было сказано в § 4.13, цепь эмиттера транзистора можно представить в виде параллельного соединения активного сопротивления эмиттера, диффузионной и барьерной емкостей (рис. 4.36). При этом часть тока эмиттера, проходящая через активное сопротивление и диффузионную емкость эмиттера, связана с инжекцией носителей заряда в базу, а часть тока, проходящая через барьерную емкость, не связана с инжекцией.

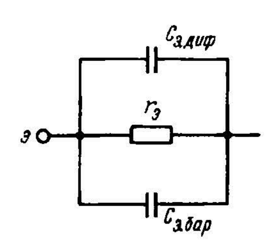


Рис. 4.36. Эквивалентная схема эмиттерной цепи транзистора

Согласно определению, данному в § 4.12, эффективность эмиттера на переменном сигнале является отношением тока, связанного с инжекцией, к полному току:

$$\dot{\gamma} = \frac{i_{э.инж}}{i_{э.инж} + i_{э.емк}} = \frac{\frac{1}{r_э} + j\omega C_{э.диф}}{\frac{1}{r_э} + j\omega C_{э.диф} + j\omega C_{э.бар}}. \qquad (4.124)$$

Как сказано в § 4.12, постоянная времени

$$r_э C_{э.диф} \approx t_{прол} \qquad (4.125)$$

Таким образом, вплоть до частот ω, сравнимых с $1/t_{прол}$, величину $\omega r_э C_{э.диф}$ можно считать малой по сравнению с единицей. Тогда окончательно

$$\dot{\gamma} \approx \frac{1}{1 + j\omega r_э C_{э.бар}}. \qquad (4.126)$$

Полученное выражение свидетельствует о том, что с увеличением частоты из-за шунтирующего действия барьерной емкости эмиттера доля инжектированных носителей заряда уменьшается. Кроме того, появляется фазовый сдвиг между общим током эмиттера и током инжектируемых в базу носителей. Так как с ростом тока сопротивление эмиттера падает быстрее, чем растет емкость, постоянная времени цепи эмиттера с увеличением постоянного тока уменьшается.

Пролет носителей заряда через базу. Точный учет искажения сигнала при пролете носителей через базу довольно сложен. Однако можно сказать, что ток коллектора должен отставать по фазе от тока эмиттера из-за конечного времени пролета носителей. Кроме того, и амплитуда тока коллектора должна быть меньше амплитуды тока эмиттера из-за размытия пакетов носителей заряда (см. § 4.9). Такая зависимость коэффициента переноса обычно аппроксимируется формулой

$$\dot{\alpha}_п = \frac{\alpha_п \exp\left(-jm_п \dfrac{\omega}{\omega_{\alpha_п}}\right)}{1 + j\dfrac{\omega}{\omega_{\alpha_п}}}. \qquad (4.127)$$

Из (4.127) модуль коэффициента переноса

$$|\dot{\alpha}_п| = \frac{\alpha_п}{\sqrt{1 + \left(\dfrac{\omega}{\omega_{\alpha_п}}\right)^2}}, \qquad (4.128)$$

а аргумент

$$\arg \dot{\alpha}_п = -\left(m_п \frac{\omega}{\omega_{\alpha_п}} + \operatorname{arctg}\frac{\omega}{\omega_{\alpha_п}}\right). \qquad (4.129)$$

Из выражения (4.128) следует, что $\omega_{\alpha_п}$ — частота, на которой модуль коэффициента переноса уменьшается в $\sqrt{2}$ раз по сравнению с низкочастотным значением. Значения коэффициента $m_п$ зависят от распределения примесей в базе и обычно лежат в пределах от 0,22 для бездрейфовых до 0,6—0,9 для дрейфовых транзисторов.

Более простая аппроксимация частотной зависимости коэффициента переноса имеет вид

$$\dot{\alpha}_п = \frac{1}{1 + j\omega t_{прол}}, \qquad (4.130)$$

где $t_{прол}$ — время пролета носителей заряда через базу.

Как показывают расчеты, эта аппроксимация дает несколько заниженные значения модуля коэффициента переноса, а на высоких частотах — и фазового сдвига. Но зато время пролета носителей через базу легко найти из их распределения. Действительно, плотность тока

$$J_p = qpv_p, \quad (4.131)$$

отсюда

$$v_p = dx/dt = J_p/(qp). \quad (4.132)$$

Подставив распределение концентрации носителей заряда из (4.12) и проинтегрировав от 0 до w, получим

$$t_{\text{прол}} = \int_0^w \left(\frac{1}{N} \int_0^w \frac{N}{D_p} dx\right) dx\,. \quad (4.133)$$

Для бездрейфового транзистора

$$t_{\text{прол}} = w^2/(2D_p), \quad (4.134)$$

для дрейфового

$$t_{\text{прол}} = \frac{w^2}{D_p \ln \frac{N(0)}{N(w)}} \left[1 - \frac{1}{\ln \frac{N(0)}{N(w)}} \left(1 - \frac{N(w)}{N(0)}\right)\right]. \quad (4.135)$$

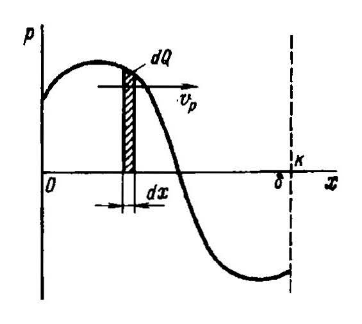


Рис. 4.37. Распределение носителей заряда в коллекторном переходе

Пролет носителей через область объемного заряда коллекторного перехода. Как сказано в § 4.9, при движении носителей через область объемного заряда коллекторного перехода в цепи коллектора проходит ток, связанный с изменением электрического поля в области перехода. Если через транзистор проходит синусоидальный сигнал, то и концентрация носителей во времени изменяется синусоидально. Кроме того, носители направленно движутся в электрическом поле перехода. Следовательно, распределение переменной составляющей концентрации носителей заряда в p-n-переходе коллектора можно представить в виде бегущей волны:

$$p = p_m \cos\omega(t - x/v_p). \quad (4.136)$$

Движущийся заряд создает наведенный ток. Если в переходе выделить слой толщиной dx (рис. 4.37), то из непрерывности тока следует, что наводимый при движении зарядов этого слоя dQ ток равен току переноса носителей в этом слое, т. е.

$$di = \frac{dQ}{a} v_p = \frac{Sqpv_p}{a} dx, \quad (4.137)$$

где a — коэффициент пропорциональности.

Токи, наводимые всеми зарядами в переходе, складываются; отсюда полный ток через коллекторный переход

$$i = \int_0^{\delta} di = \frac{Sqp_m}{a} \int_0^{\delta} v_p \cos\omega\left(t - \frac{x}{v_p}\right) dx. \qquad (4.138)$$

Положим, что скорость движения носителей v_p постоянна. Такое предположение основано на том, что напряженность электрического поля в коллекторном переходе часто превосходит ту, при которой наблюдается насыщение скорости движения носителей, так что скорость изменяется только в небольших областях у краев перехода.

Тогда, взяв интеграл в (4.138), получим

$$i = \frac{qp_m S}{a} \frac{v_p}{\omega/(2v_p)} \sin\frac{\delta\omega}{2v_p} \cos\left(\omega t - \frac{\delta\omega}{2v_p}\right). \qquad (4.139)$$

Отсюда следует, что ток через *p-n*-переход с изменением частоты изменяется как по амплитуде, так и по фазе в отличие от случая, когда эффекты, связанные с пролетом, отсутствуют. В частности, при $\omega \to 0$

$$i_0 = \frac{qSp_m v_p \delta}{a} = qSp_m v_p\,, \qquad (4.140)$$

т. е. $\delta = a$, а при других частотах

$$i = i_0 \frac{2v_p}{\delta\omega} \sin\frac{\delta\omega}{2v_p} \cos\left(\omega t - \frac{\delta\omega}{2v_p}\right). \qquad (4.141)$$

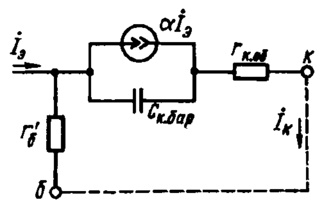


Рис. 4.38. Эквивалентная схема коллекторной цепи транзистора

Амплитуда тока коллектора падает в $\sqrt{2}$ раз по сравнению с низкочастотным значением, когда $\sin\frac{\delta\omega}{2v_p} = \frac{1}{\sqrt{2}}\,\frac{\delta\omega}{2v_p}$, что соответствует $\frac{\delta\omega}{2v_p} = 1{,}39$. По аналогии с (4.130) выражение (4.141) можно переписать в виде

$$\dot\alpha^{*} = \frac{i}{i_0} = \frac{1}{1 + j\omega 0{,}5 t_{прол.к}}, \qquad (4.142)$$

где $t_{прол.к} = \delta / v_p$.

Постоянная времени цепи коллектора. Коллектор реального транзистора обладает некоторым сопротивлением, особенно заметным у транзисторов, изготовляемых методом диффузии. Если рассматривать область только высоких частот, то эквивалентную схему рис. 4.31, *б* можно преобразовать в схему, показанную на рис. 4.38. При этом нужно учесть, что из-за сравнительно большого значения диффузионной емкости базы она на высоких частотах обладает малым сопротивлением, так что все сопротивление цепи базы только объемное. В режиме короткого замыкания цепи коллектора емкость $C_{кб}$ не играет роли, а емкость $C_{кэ}$ будем считать малой.

Ток коллектора в режиме короткого замыкания определим из соотношения

$$I_к\left(r_{к.об} + r'_б + \frac{1}{j\omega C_{к.бар}}\right) = I_э\left(\frac{\dot\alpha}{j\omega C_{к.бар}} + r'_б\right). \qquad (4.143)$$

или

$$\dot I_к = \dot I_э \frac{\dfrac{\dot\alpha}{j\omega C_{к.бар}} + r'_б}{r_{к.об} + r'_б + \dfrac{1}{j\omega C_{к.бар}}}, \qquad (4.144)$$

откуда

$$-\dot h_{21б} = \frac{\dot I_к}{\dot I_э} = \frac{\dot\alpha + j\omega r'_б C_{к.бар}}{1 + j\omega r'_б C_{к.бар} + j\omega r_{к.об} C_{к.бар}}. \qquad (4.145)$$

Следовательно, из-за влияния постоянной времени цепи коллектора коэффициент передачи тока транзистора также может изменяться как по амплитуде, так и по фазе.

Чтобы упростить расчеты и выделить влияние постоянной времени цепи коллектора, положим, что все остальные факторы, определяющие частотные свойства транзистора, несущественны. В этом случае $\dot\alpha = 1$. Тогда

$$-\dot h_{21б} = \frac{1}{1 + j\omega r_{к.об} C_{к.бар} \dfrac{1}{1 + j\omega r'_б C_{к.бар}}} \approx \frac{1}{1 + j\omega r_{к.об} C_{к.бар}} = \dot\alpha_2^* \qquad (4.146)$$

Коэффициент передачи тока реального транзистора. Учесть одновременно все факторы, влияющие на частотные свойства реального транзистора, крайне сложно. Для оценки частотных свойств можно считать, что полный коэффициент передачи тока транзистора равен произведению коэффициентов передачи, определяемых процессами в отдельных областях. Тогда, аппроксимировав $\dot h_{21б}$ выражением, аналогичным (4.78), получим

$$-\dot h_{21б} = \dot\gamma \dot\alpha_п \dot\alpha_1^* \dot\alpha_2^* . \qquad (4.147)$$

Частотную зависимость $\dot h_{21б}$ можно аппроксимировать выражением

$$-\dot h_{21б} = \frac{h_{21б}}{1 + j\dfrac{\omega}{\omega_{гр}}} \approx \frac{1}{1 + j\dfrac{\omega}{\omega_{гр}}}. \qquad (4.148)$$

Здесь $\omega_{гр}$ — частота (физический смысл будет уточнен позже), значение которой характеризует частотные свойства транзистора; на частотах около $\omega_{гр}$ заметно падает $|\dot h_{21б}|$ и возрастает фазовый сдвиг между токами коллектора и эмиттера. Подставив значения $\dot\gamma$, $\dot\alpha_п$, $\dot\alpha_1^*$, $\dot\alpha_2^*$ (полученные ранее), найдем

$$\frac{1}{1 + j\dfrac{\omega}{\omega_{гр}}} = \frac{1}{1 + j\omega r_э C_{э.бар}} \frac{1}{1 + j\omega t_{прол}} \times$$

$$\times \frac{1}{1 + j\omega \dfrac{t_{прол.к}}{2}} \frac{1}{1 + j\omega r_{к.об} C_{к.бар}}. \qquad (4.149)$$

Даже такое упрощенное выражение довольио сложио для анализа. Например, для расчета частоты $\omega_{гр}$ по этому соотношению пришлось бы решить уравнеиие восьмой степени. Поэтому ограничимся приближеииым рассмотреиием, заключающимся в том, что каждая постояииая времени в (4.149) считается ие очеиь большой.

Тогда

$$\frac{1}{1+j\frac{\omega}{\omega_{гр}}} \approx \frac{1}{1+j\omega r_э C_{э.бар} + j\omega t_{прол} + j\omega \frac{t_{прол.к}}{2} + j\omega r_{к.об} C_{к.бар}}, \quad (4.150)$$

откуда

$$\frac{1}{\omega_{гр}} \approx r_э C_{э.бар} + t_{прол} + \frac{t_{прол.к}}{2} + r_{к.об} C_{к.бар}\,. \quad (4.151)$$

Таким образом, полная постояииая временн транзистора в первом приближении равна сумме постояниых времен, определяемых различными процессами.

Влияние различных процессов на частотные свойства реального траизистора зависит от его конструкции и режима работы. Так, для бездрейфовых транзисторов характерио, что время пролета иосителей через базу преобладает иад всеми другими постояниыми времеии.

Следовательио, для них

$$1/\omega_{гр} \approx t_{прол} = w^2/(2D_p)\,. \quad (4.152)$$

Для дрейфовых транзисторов существенную роль играет постояиная времени цепи эмиттера, особенно при малых токах.

Для сверхвысокочастотных транзисторов все факторы могут оказаться примерио равнозиачиыми.

Зависимость частоты $\omega_{гр}$ от тока эмиттера. Наличие нескольких факторов, влияющих на частотные свойства, обусловливает довольно сложную токовую зависимость $\omega_{гр}$ (рис. 4.39). В области малых токов определяющую роль играет $r_э C_{э.бар}$. С увеличением тока сопротивление эмиттера падает и определяющим процессом становится перенос носителей через базу. При этом возникновение электрического поля в базе приводит к дальнейшему росту $\omega_{гр}$. При очень больших токах иачинает сказываться сдвиг граиицы области объемного заряда коллекториого перехода и увеличение его толщины (см. § 4.5). Это ведет к снижению $\omega_{гр}$.

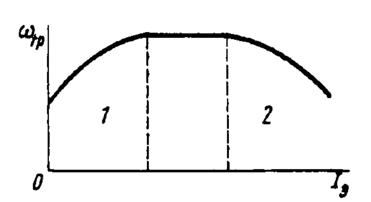


Рис. 4.39. Зависимость граничной частоты коэффициента передачи тока базы траизистора от постоянного тока эмнттера:

1 — диапазон частот с преимущественным влиянием цепи эмиттера; 2 — диапазон частот с преимущественным влиянием цепи коллектора из-за увеличения толщины коллекторного перехода

Уточненная частотная характеристика. В некоторых случаях для аппроксимации частотной характеристики транзистора используют выражение, аналогичное (4.127):

$$\dot{h}_{21б} = \frac{h_{21б} \exp\left(-jm\frac{\omega}{\omega_{h21б}}\right)}{1 + j\frac{\omega}{\omega_{h21б}}}. \qquad (4.153)$$

Как и ранее, зависимость модуля $\dot{h}_{21б}$ от частоты имеет вид

$$|\dot{h}_{21б}| = \frac{h_{21б}}{\sqrt{1 + \left(\frac{\omega}{\omega_{h21б}}\right)^2}}, \qquad (4.154)$$

а аргумента

$$\arg \dot{h}_{21б} = -\left(m\frac{\omega}{\omega_{h21б}} + \operatorname{arctg}\frac{\omega}{\omega_{h21б}}\right). \qquad (4.155)$$

Физический смысл *предельной частоты коэффициента передачи тока эмиттера* $\omega_{h21б}$ — частота, на которой модуль коэффициента передачи тока эмиттера падает в $\sqrt{2}$ раз по сравнению с его низкочастотным значением.

Из равенства аргументов выражений (4.155) и (4.148) получим

$$m\frac{\omega}{\omega_{h21б}} + \operatorname{arctg}\frac{\omega}{\omega_{h21б}} = \operatorname{arctg}\frac{\omega}{\omega_{гр}} \qquad (4.156)$$

или, взяв приближенно

$$\operatorname{arctg}\frac{\omega}{\omega_{h21б}} \approx \frac{\omega}{\omega_{h21б}} \text{ и } \operatorname{arctg}\frac{\omega}{\omega_{гр}} \approx \frac{\omega}{\omega_{гр}}, \qquad (4.157)$$

$$\omega_{h21б} = (1 + m)\omega_{гр}. \qquad (4.158)$$

Частотная зависимость $\dot{h}_{21э}$. Воспользовавшись выражением (4.153), найдем частотную зависимость коэффициента передачи тока для схемы с общим эмиттером $\dot{h}_{21э}$. Из условия баланса токов в транзисторе

$$\dot{h}_{21э} = \frac{-\dot{h}_{21б}}{1 + \dot{h}_{21б}}. \qquad (4.159)$$

Подставив значение $\dot{h}_{21б}$ из (4.148), получим

$$\dot{h}_{21э} = \frac{-h_{21б}}{\left(1 + j\frac{\omega}{\omega_{гр}}\right)\left(1 + \frac{h_{21б}}{1 + j\frac{\omega}{\omega_{гр}}}\right)} = \frac{-h_{21б}}{1 + h_{21б} + j\frac{\omega}{\omega_{гр}}}, \qquad (4.160)$$

или

$$\dot{h}_{21э} = \frac{h_{21э}}{1 + j\frac{\omega}{\omega_{h21э}}}, \qquad (4.161)$$

где

$$\omega_{h21э} = \omega_{гр}(1 - |h_{21б}|) \approx \omega_{гр}/h_{21э} \,. \qquad (4.162)$$

Физический смысл *предельной частоты коэффициента передачи тока базы* $\omega_{h21э}$ — частота, на которой модуль коэффициента передачи тока базы уменьшается в $\sqrt{2}$ раз по сравнению с низкочастотным значением.

Следует отметить, что если частотные свойства транзистора определяются только процессами переноса через базу (например, для бездрейфовых транзисторов), то

$$\omega_{гр} = 1/t_{прол} \,; \quad h_{21э} = \tau_p/t_{прол} \qquad (4.163)$$

и

$$\omega_{h21э} = 1/\tau_p \,. \qquad (4.164)$$

Практически частота $\omega_{h21э}$ сравнительно невелика — транзистор может работать и на существенно более высоких частотах. В этом случае можно считать, что

$$|\dot{h}_{21э}| \approx \left| \frac{h_{21э}\,\omega_{h21э}}{\omega} \right| = \frac{\omega_{гр}}{\omega} \,. \qquad (4.165)$$

Таким образом, частота $\omega_{гр}$ — *граничная частота коэффициента передачи тока базы* — это частота, при которой модуль коэффициента передачи тока базы становится равным единице.

Из соотношения (4.165) следует, что на частотах, превышающих предельную частоту коэффициента передачи тока базы $\omega_{h21э}$, модуль коэффициента передачи тока базы обратно пропорционален частоте. Поэтому для экспериментального определения граничной частоты коэффициента передачи тока базы $\omega_{гр}$ достаточно измерить модуль коэффициента передачи тока базы $|h_{21э}|$ на какой-либо частоте $\omega > \omega_{h21э}$ и далее использовать соотношение

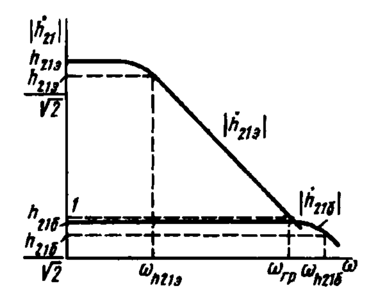


Рис. 4.40. Зависимости модулей коэффициентов передачи от частоты (масштаб логарифмический)

$$\omega_{гр} = |\dot{h}_{21э}|\,\omega \,. \qquad (4.166)$$

Графически зависимость коэффициентов передачи от частоты представлена на рис. 4.40. Из графика видно, что $|\dot{h}_{21э}|$ падает на значительно более низких частотах, чем $|\dot{h}_{21б}|$. Физически это связано с влиянием фазового сдвига между токами эмиттера и коллектора (рис. 4.41). С увеличением частоты фазовый сдвиг растет, и это приводит к росту тока базы, даже если ток

коллектора и ток эмиттера неизменны по абсолютному значению. Следовательно, и

$$|\dot{h}_{21э}| = \dot{I}_к/\dot{I}_б \tag{4.167}$$

уменьшается.

Усиление по мощности на высоких частотах. С ростом частоты сигнала коэффициент усиления транзистора по мощности падает. Это связано как с уменьшением коэффициента усиления по току, рассмотренным ранее, так и с влиянием цепи $r'_б C_{к.бар}$.

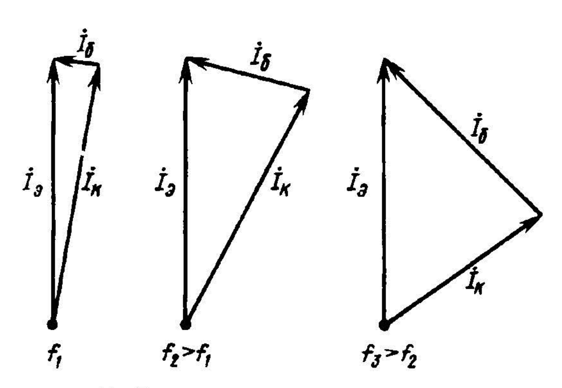


Рис. 4.41. Векторные диаграммы токов транзистора на разных частотах

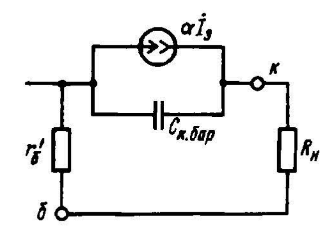


Рис. 4.42. Коллекторная цепь транзистора (для переменной составляющей тока)

С ростом частоты шунтирующее действие этой цепи на нагрузку усиливается, в нагрузку ответвляется меньший ток, что и приводит к снижению коэффициента усиления по мощности (рис. 4.42).

Для борьбы с этим явлением следует выбирать определенное сопротивление нагрузки, т. е. согласовывать нагрузку. Однако на очень высоких частотах и это не дает возможности получить достаточный коэффициент усиления по мощности. На некоторой частоте коэффициент усиления транзистора по мощности даже при согласованной нагрузке становится равным единице. Это означает, что транзистор на такой частоте уже нельзя рассматривать как активный элемент электрической схемы.

При коэффициенте усиления по мощности, равном (или меньшем) единице, нельзя осуществить режим самовозбуждения в генераторе на транзисторе. Поэтому ту частоту, на которой коэффициент усиления по мощности становится равным единице, называют *максимальной частотой генерации*.

Расчет показывает, что максимальную частоту генерации можно найти по формуле

$$f_{max} = \sqrt{\frac{h_{21б} f_{гр}}{8\pi r'_б C_{к.бар}}}. \qquad (4.168)$$

Максимальная частота генерации наиболее удачно характеризует частотные свойства транзистора, так как определяет диапазон частот, в котором транзистор остается активным элементом электрической схемы.

§ 4.15. РАБОТА ТРАНЗИСТОРА НА ИМПУЛЬСАХ

Особенности работы

Транзисторы часто применяют в импульсных устройствах и в качестве транзисторного ключа. При работе транзистора в импульсных устройствах от него, как правило, требуется неискаженное воспроизведение усиленного импульса на выходе. Работа транзистора при усилении малых импульсных сигналов в принципе ничем не отличается от работы транзистора при усилении малых синусоидальных сигналов. Импульс можно представить в виде суммы ряда гармонических составляющих и, зная частотные свойства транзистора, определить искажения формы импульса, которые могут иметь место при усилении. Работа транзистора при усилении больших импульсных сигналов отличается тем, что транзистор в этом случае может оказаться не только в активном режиме, но и в режимах отсечки и насыщения.

При работе транзистора в качестве транзисторного ключа необходимо, чтобы сопротивление транзистора на выходе, т. е. в цепи нагрузки, резко изменялось под влиянием входного управляющего импульса. Для этого амплитуда входных импульсов должна быть достаточной для перевода транзистора из режима отсечки в активный режим работы и далее в режим насыщения, а также в обратном направлении.

Схема с общей базой

Рассмотрим процессы, происходящие в транзисторе, включенном по схеме с общей базой, при прохождении через эмиттер импульса тока длительностью $t_{имп}$ в прямом направлении с последующим изменением его направления на обратное (рис. 4.43). В исходном состоянии транзистор находится в режиме отсечки, т. е. эмиттерный и коллекторный переходы смещены в обратном направлении.

После подачи через эмиттер импульса тока в прямом направлении ток коллектора появляется не сразу, так как необходимо некоторое время на перезаряд барьерных емкостей эмиттерного и коллекторного переходов, а также на передвижение инжектированных неосновных носителей заряда до коллекторного пере-

хода (рис. 4.43, *б*). Интервал времени между моментом подачи на вход транзистора импульса тока и моментом достижения выходным током значения, соответствующего 10% его амплитуды, называют *временем задержки для биополярного транзистора* $t_{зд}$.

В дальнейшем продолжается процесс перезарядки барьерной емкости эмиттерного перехода, что приводит к увеличению напряжения на эмиттерном переходе, к увеличению граничной концентрации и градиента концентрации неосновных носителей заряда в базе около эмиттерного перехода (рис. 4.43, *в*). Увеличение градиента концентрации неосновных носителей заряда около эмиттерного перехода соответствует увеличению инжек-

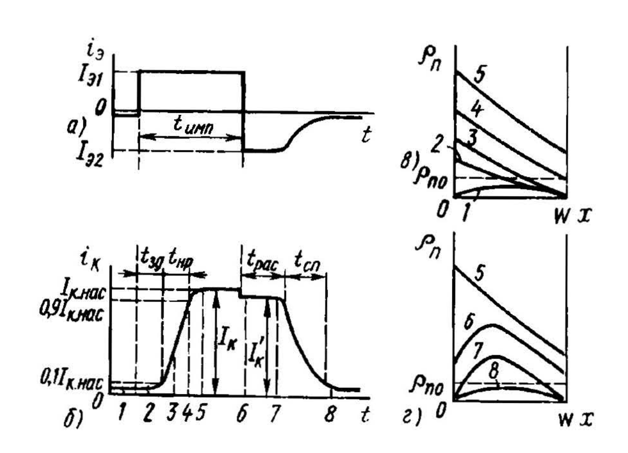


Рис. 4.43. Временные зависимости тока эмиттера (*а*) и тока коллектора (*б*) при работе транзистора в качестве ключа по схеме с общей базой и распределение неосновных носителей заряда в базе в различные моменты времени (*в*, *г*)

ционной составляющей тока эмиттера. Емкостная же составляющая тока эмиттера по мере заряда барьерной емкости эмиттерного перехода уменьшается, так что полный ток эмиттера $I_{э1}$ остается неизменным (рис. 4.43, *а*). Его значение определяется параметрами генератора тока во входной цепи транзистора.

В связи с увеличением инжекционной составляющей тока эмиттера происходит процесс накопления неосновных носителей заряда в базе транзистора. Этот процесс также происходит не мгновенно, так как скорость движения неосновных носителей заряда в базе конечна.

В процессе накопления неосновных носителей заряда увеличивается градиент их концентрации около коллекторного перехода, что соответствует увеличению тока коллектора. При больших значениях тока эмиттера $I_{э1}$ ток коллектора ограничен не током эмиттера, а параметрами выходной коллекторной цепи. Эмиттер инжектирует в базу такое количество неосновных носителей заряда, которое коллекторный переход не может экстра-

гировать при заданном значении сопротивления нагрузки и ЭДС источника питания в цепи коллектора. Поэтому в базе транзистора около коллекторного перехода начинает расти граничная концентрация неосновных носителей заряда. Когда эта граничная концентрация неосновных носителей заряда превысит значение равновесной концентрации неосновных носителей заряда, транзистор перейдет из активного режима в режим насыщения. В этот момент (кривая *4* на рис. 4.43, *в*) ток коллектора

$$I_{\text{К нас}} \approx \mathcal{E}_{\text{КБ}} / R_{\text{н}} . \qquad (4.169)$$

В действительности значение установившегося тока коллектора транзистора, находящегося в режиме насыщения, несколько превышает значение тока насыщения, вычисленного по (4.169), так как кроме ЭДС источника питания надо еще учесть падение напряжения на объемном сопротивлении базы. При прохождении через эмиттер тока в прямом направлении падение напряжения на объемном сопротивлении базы, как это видно на рис. 4.44, должно складываться с ЭДС источника питания в коллекторной цепи:

$$I_{\text{К}} = I_{\text{К нас}} = (\mathcal{E}_{\text{КБ}} + \Delta U_1) / R_{\text{н}} . \qquad (4.170)$$

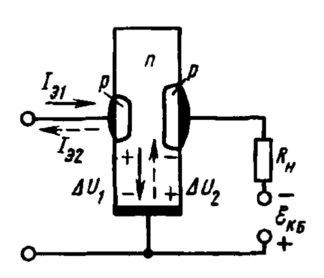


Рис. 4.44. Пояснение скачкообразного изменения тока коллектора из-за изменения полярности падения напряжения на объемном сопротивлении базы при изменении направления тока эмиттера

Интервал времени, в течение которого ток коллектора нарастает от 10 до 90% его амлитуды, называют *временем нарастания для биполярного транзистора* $t_{\text{нр}}$ (см. рис. 4.43, *б*). Интервал времени, являющийся суммой времени задержки и времени нарастания, называют *временем включения биполярного транзистора* $t_{\text{вкл}}$. Время включения биполярного транзистора зависит от амплитуды импульса прямого тока эмиттера и от частотных свойств транзистора, а также от $\mathcal{E}_{\text{КБ}}$ и $R_{\text{н}}$.

В момент изменения направления тока эмиттера происходит изменение полярности падения напряжения на объемном сопротивлении базы (рис. 4.44). При этом скачкообразно уменьшается значение тока коллектора, так как

$$I'_{\text{К}} = (\mathcal{E}_{\text{КБ}} - \Delta U_2) / R_{\text{н}} . \qquad (4.171)$$

Одновременно начинается процесс рассасывания неосновных носителей заряда, накопленных в базе транзистора. В первый момент после изменения направления тока эмиттера граничные концентрации неосновных носителей в базе около *p-n*-переходов эмиттера и коллектора велики: они превышают значение равновесной концентрации неосновных носителей. Поэтому сопротивления этих переходов для обратных токов оказываются

очень малыми. Значения обратных токов эмиттера и коллектора определяются в основном сопротивлениями во виешних цепях и ЭДС источииков питаиия. Граничиые концентрации иеосиовных носителей в базе около p-n-переходов ие могут мгновеино после переключения входа транзистора умеиьшиться до нуля. Это соответствовало бы бескоиечио большим градиеитам концеитрации неосновных иосителей заряда в базе около p-n-переходов и бескоиечно большим токам, чего практически быть не может из-за коиечных зиачеиий сопротивления во внешиих цепях траизистора. До тех пор пока в процессе рассасываиия граничиые коицеитрации иеосиовиых иосителей заряда в базе около p-n-переходов ие умеиьшатся до иуля, обратиые токи через соответствующие p-n-переходы будут оставаться постоянными, т. е. токи эмиттера и коллектора будут неизмениыми, пока траизистор будет иаходиться в режиме насыщения.

После умеиьшеиия граничиых коицеитраций неосиовных носителей заряда в базе около переходов до иуля будут уменьшаться со времеием токи эмиттера и коллектора, так как процесс рассасываиия иеосиовиых носителей заряда из базы продолжается и умеиьшается абсолютиое значение градиентов коицеитрации иеосновиых иосителей заряда около соответствующих p-n-переходов. Измеиения в распределеиии иеосиовиых иосителей заряда в различные моменты времеии процесса рассасываиия показаиы иа рис. 4.43, *г*.

Интервал времеии между момеитом подачи иа вход траизистора запирающего импульса и моментом, когда ток коллектора достигиет задаиного уровня (иапример $0{,}9 I_{К\,нас}$), называют *временем рассасывания для биполярного транзистора* $t_{рас}$. Интервал времеии между момеитами спада выходного тока от зиачения, соответствующего 90% его амплитуды, до зиачения, соответствующего 10% его амплитуды, называют *временем спада для биполярного транзистора* $t_{сп}$ (см. рис. 4.43, *б*). Интервал времеии между момеитом подачи на вход траизистора запирающего импульса и момеитом, когда ток коллектора достигиет зиачеиия, соответствующего 10% его амплитудиого значения, иазывают *временем выключения биполярного транзистора* $t_{выкл}$. Время рассасываиия, время спада и соответствеино время выключеиия траизистора зависят от амплитуды импульса прямого включающего тока эмиттера, от ЭДС источиика питаиия и сопротивлеиия нагрузки в коллекторной цепи, а также от частотиых свойств транзистора.

Увеличить быстродействие траизистора, работающего в качестве электроииого ключа, т. е. умеиьшить время рассасываиия, можио путем введеиия в кристалл полупроводиика примесей рекомбииациониых ловушек (золото для кремиия). При этом будет умеиьшеио время жизии иеосиовиых иосителей заряда. Одиако наряду с увеличеиием быстродействия в таких транзисторах, во-первых, будут меиьше коэффициеиты передачи тока из-за более интенсивной рекомбииации иеосновиых носителей заряда в базе.

Во-вторых, у них будет больше обратный ток коллектора и эмиттера из-за более интенсивной тепловой генерации носителей заряда в коллекторном и эмиттерном *p-n*-переходах, а также в прилегающих к этим переходам областях.

Более удачным методом повышения быстродействия транзистора, работающего в качестве электронного ключа, является шунтирование коллекторного перехода диодом Шотки, в котором при прямом смещении отсутствуют инжекция неосновных носителей заряда и их накопление. Структура и принцип действия такого транзистора с диодом Шотки будут рассмотрены в § 7.4, так как наибольшее распространение такие транзисторы получили в интегральных микросхемах.

Схема с общим эмиттером

В транзисторе, включенном по схеме с общим эмиттером, при работе на импульсах с большой амплитудой происходят те же процессы накопления и рассасывания неосновных носителей заряда в базе. На рис. 4.45 показаны временные зависимости тока базы и тока коллектора при включении транзистора по схеме с общим эмиттером. Особенностью временной зависимости в этом случае по сравнению с аналогичной зависимостью тока коллектора в схеме с общей базой является скачкообразное увеличение тока коллектора при перемене направления тока базы с $I_{Б1}$ до $I_{Б2}$. Входному включающему току $I_{Б1}$ соответствует отрицательный потенциал базового вывода $U_{Б1}$ по отношению к общему эмиттерному выводу (рис. 4.46). Поэтому пренебрегая сопротивлением транзистора в режиме насыщения,

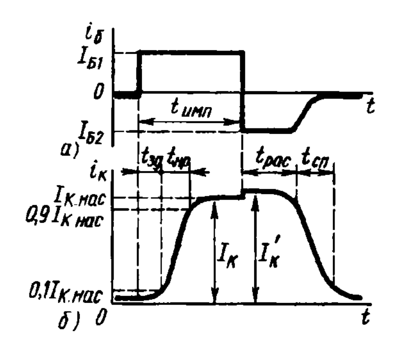


Рис. 4.45. Временные зависимости тока базы (*а*) и тока коллектора (*б*) при работе транзистора в качестве ключа по схеме с общим эмиттером

$$I_К \approx (\mathscr{E}_{КЭ} - U_{Б1})/R_н.$$

Выключающему току базы $I_{Б2}$ соответствует положительный потенциал базового вывода $U_{Б2}$, поэтому, также пренебрегая сопротивлением транзистора, который еще находится в режиме насыщения, получаем

$$I'_К \approx (\mathscr{E}_{КЭ} + U_{Б2})/R_н.$$

Изменение тока коллектора в момент переключения входа транзистора в схеме с общим эмиттером обычно небольшое по сравнению с изменением того же тока в схеме с общей базой.

Качество работы транзистора в схеме электронного ключа оценивают не только по параметрам, характеризующим инер-

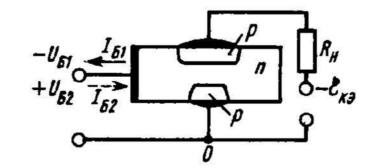


Рис. 4.46. Пояснение скачкообразного изменения тока коллектора при перемене полярности входного напряжения

ционность процессов в транзисторе при его переключении (время задержки, время нарастания, время рассасывания, время спада), но и по параметрам, характеризующим выходное и входное сопротивления транзистора в режиме насыщения. Важнейшим из них является напряжение насыщения коллектор — эмиттер $U_{КЭнас}$ — напряжение между выводами коллектора и эмиттера транзистора в режиме насыщения при заданных токах базы и коллектора.

§ 4.16. ШУМЫ В ТРАНЗИСТОРАХ

Транзистор, как и другие полупроводниковые приборы, обладает собственными шумами, т. е. беспорядочными колебаниями тока и напряжения на выходе при отсутствии сигналов на входе.

Физические явления, обусловливающие появление шумов в транзисторе, по существу, те же, что и в других полупроводниковых приборах, т. е. шумы в транзисторе можно подразделить на тепловые, дробовые и избыточные.

Тепловые шумы вызваны беспорядочным перемещением зарядов в полупроводнике из-за теплового хаотического движения носителей заряда, что сопровождается флуктуациями тока или электродвижущей силы, определяемой по формуле американского физика Х. Найквиста:

$$\mathcal{E}^2 = 4kTR\Delta f, \tag{4.172}$$

где Δf — полоса частот, внутри которой измеряют флуктуации ЭДС.

Скорости теплового хаотического движения носителей заряда обычно значительно превышают скорости дрейфа этих носителей в электрическом поле. Поэтому интенсивность тепловых шумов не зависит ни от приложенного напряжения, ни от тока, ни от частоты (а только от полосы частот, в которой происходит измерение шумов).

Дробовые шумы связаны с дискретностью заряда носителей и со случайным характером инжекции и экстракции этих носителей через электрические переходы. Среднеквадратичное значение дробового шума

$$I^2 = 2qI_Э\Delta f. \tag{4.173}$$

Избыточные шумы вызваны неравномерностью процессов генерации и рекомбинации носителей заряда, а также процессами захвата и освобождения носителей заряда ловушками захвата. Все эти процессы наиболее интенсивно происходят на поверхности полупроводника. Спектральный состав этих шумов

связан с временем жизни носителей заряда. Поэтому интенсивность избыточных шумов уменьшается с ростом частоты примерно обратно пропорционально частоте.

Основным параметром, характеризующим шумы в транзисторах, считают *коэффициент шума*

$$F = \frac{(P_{сигн}/P_{ш})_{вх}}{(P_{сигн}/P_{ш})_{вых}}, \tag{4.174}$$

где $P_{сигн}$ — мощность сигнала; $P_{ш}$ — мощность шумов.

Коэффициент шума показывает, во сколько раз ухудшается отношение сигнал/шум из-за наличия собственных шумов усилителя (в данном случае транзистора). Для идеального усилителя $F = 1$. Иногда коэффициент шума выражают в децибелах. Тогда

$$F = 10 \lg \frac{(P_{сигн}/P_{ш})_{вх}}{(P_{сигн}/P_{ш})_{вых}}. \tag{4.175}$$

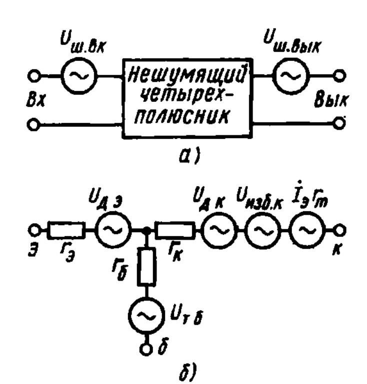


Рис. 4.47. Шумовые эквивалентные схемы транзисторов:
а — формальная; *б* — физическая

При описании шумовых свойств транзистора можно использовать его шумовую эквивалентную схему, принцип построения которой примерно тот же, что и обычных эквивалентных схем. При построении формальной эквивалентной схемы транзистор можно рассматривать как идеальный нешумящий четырехполюсник, во входную и выходную цепи которого включены генераторы шумов (рис. 4.47, *а*). Этот четырехполюсник можно заменить эквивалентной схемой, например как в § 4.11. Значения ЭДС (или тока) генераторов шумов можно определить экспериментально.

Физические шумовые эквивалентные схемы дают возможность связать источники шумов в транзисторе с коэффициентом шума. При этом генераторы шумов, включаемые в эквивалентную схему, связывают с определенными процессами (рис. 4.47, *б*). Шумы эмиттерного перехода транзистора являются дробовыми и отражаются соответствующим генератором, шумы активного сопротивления базы — тепловыми. На коллекторном переходе имеются и дробовые, и избыточные шумы. ЭДС генераторов шумов можно определить следующим образом.

Через эмиттерный переход идет ток; значит, там должны быть дробовые шумы. В соответствии с формулой (4.173) ЭДС от этих шумов, выделяемая на сопротивлении эмиттера (для единичной полосы частот),

$$\mathcal{E}_{дэ} = \sqrt{2qI_э r_э^2}. \tag{4.176}$$

Аналогично, ЭДС дробовых шумов в цепи коллектора

$$\mathcal{E}_{дк}=\sqrt{2qI_К r_к^2}. \qquad (4.177)$$

База представляет собой активное сопротивление; следовательно, здесь должны наблюдаться тепловые шумы с ЭДС

$$\mathcal{E}_{тб}=\sqrt{4kTr'_б}. \qquad (4.178)$$

ЭДС избыточных шумов отнесенных полностью к цепи коллектора, можно определить по эмпирической формуле

$$\mathcal{E}_{изб.к}=KU_{КБ}^{a}\left(\frac{U_{КБ}}{I_К}\right)^{в}\frac{1}{f}, \qquad (4.179)$$

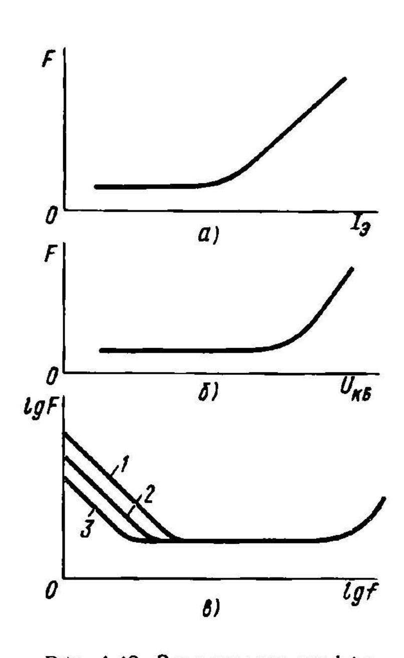


Рис. 4.48. Зависимости коэффициента шума транзистора:

а — от тока эмиттера; *б* — от напряжения на коллекторе; *в* — от частоты (*1* — загрязненная поверхность кристалла; *2* — нормально обработанная поверхность; *3* — поверхность с улучшенной обработкой)

где K — коэффициент, зависящий от материала и обработки поверхности полупроводника; a и b — показатели степени, имеющие значения в пределах 1,2—1,8.

Пользуясь эквивалентной схемой (рис. 4.47, *б*), можно определить шумы, создаваемые транзистором в схеме.

Коэффициент шума транзистора зависит от режима его работы и частоты. Зависимость коэффициента шума от тока транзистора представлена на рис. 4.48, *а*, из которого видно, что коэффициент шума при малых $I_Э$, когда преобладают тепловые и избыточные шумы, сначала сравнительно слабо зависит от тока, а при дальнейшем увеличении тока в связи с увеличением роли дробовых шумов растет примерно пропорционально $I_Э$.

С изменением напряжения коллектора коэффициент шума также сначала изменяется мало (рис. 4.48, *б*), а затем растет довольно быстро. Это объясняется тем, что при низких напряжениях коэффициент шума определяется тепловыми и дробовыми шумами, а при больших напряжениях начинают преобладать избыточные шумы.

При росте частоты (рис. 4.48, *в*) коэффициент шума сначала падает из-за снижения роли избыточных шумов, затем в некотором диапазоне частот остается постоянным. Здесь он определяется тепловыми и дробовыми шумами. На частотах, где уменьшается коэффициент усиления транзистора, коэффициент шума снова растет. Это следует из определения коэффициента шума

(4.174) — при очень высоких частотах мощность сигнала на выходе падает, в то время как мощность шумов, создаваемых в транзисторе, не изменяется.

Состояние поверхности полупроводника существенно влияет на избыточные шумы (рис. 4.48, *в*). Поэтому создание малошумящих транзисторов основано на совершенствовании обработки поверхности.

§ 4.17. ТЕХНОЛОГИЯ ИЗГОТОВЛЕНИЯ И КОНСТРУКЦИЯ БИПОЛЯРНЫХ ТРАНЗИСТОРОВ

Классификация по мощности и по частоте

В зависимости от максимально допустимой мощности рассеяния биполярные транзисторы можно подразделить на транзисторы малой мощности ($P_{max} \leqslant 0{,}3$ Вт), средней мощности ($0{,}3 < P_{max} \leqslant 1{,}5$ Вт) и большой мощности ($P_{max} > 1{,}5$ Вт); в зависимости от значения граничной частоты коэффициента передачи тока — на транзисторы низкой частоты ($f_{гр} \leqslant 3$ Мгц), средней частоты (3 МГц $< f_{гр} \leqslant$ 30 МГц), высокой частоты (30 МГц $< f_{гр} \leqslant$ 300 МГц) и сверхвысокой частоты или СВЧ-транзисторы ($f_{гр} > 300$ Мгц).

Методы формирования транзисторных структур

В начальный период развития транзисторной техники биполярные транзисторы делали только из германия методом вплавления примесей — *сплавные транзисторы*. В последующие годы, после преодоления ряда трудностей очистки монокристаллического кремния, были созданы кремниевые транзисторы. Кремний обладает большей шириной запрещенной зоны. Поэтому кремниевые транзисторы могут работать при более высоких температурах (до 125°С), имеют меньшие обратные токи коллектора и эмиттера, а также более высокие пробивные напряжения. На кремниевых монокристаллах относительно просто можно создать слой диоксида кремния, который обладает маскирующими свойствами, при диффузии легирующих примесей в кремний. Это привело при производстве кремниевых транзисторов и других кремниевых приборов к широкому использованию высокопроизводительных и точных методов планарной технологии. В связи с перечисленными преимуществами кремниевые биполярные транзисторы практически полностью вытеснили аналогичные германиевые приборы.

Основным методом формирования транзисторных структур современных транзисторов является планарная технология (см. § 2.4). Транзисторы, выполненные по этой технологии, называют *планарными*. Одним из преимуществ планарной техно-

логии является ее универсальность, позволяющая на одном и том же оборудовании организовать производство различных по параметрам транзисторов путем изменения набора фотошаблонов и режимов диффузии примесей.

При планарной технологии можно создавать транзисторы с хорошими частотными свойствами. Это обусловлено тем, что в данном случае можно проводить селективную диффузию, т. е. вводить примеси только в небольшие ограниченные области, строго контролируя глубину диффузии. В то же время оптические методы, применяемые при фотолитографии, позволяют с большей точностью совмещать эти области. В результате возможно изготовление транзисторов с толщиной базы в доли микрометра и размерами выпрямляющих электрических переходов в единицы микрометров. Граничная частота коэффициента передачи тока биполярных транзисторов достигает 10 ГГц.

Места выхода *p-n*-переходов планарного транзистора на поверхность кристалла полупроводника оказываются под слоем диоксида кремния, который является хорошим диэлектриком. Он служит защитой поверхности кремния от внешних воздействий, повышая стабильность параметров и надежность транзисторов. Для усиления защитных свойств слоя диоксида кремния сверху наносят тонкий слой легкоплавкого стекла.

Для уменьшения объемного сопротивления коллекторной области транзистора формирование транзисторной структуры производят в тонком эпитаксиальном слое с относительно малой концентрацией примесей, нанесенном на низкоомную подложку с электропроводностью того же типа. Например, при подложке и эпитаксиальном слое с электропроводностью *n*-типа, полученная структура транзистора n^+-p-n-n^+-типа имеет двухслойную коллекторную область, состоящую из высокоомной тонкой части эпитаксиального слоя и низкоомной подложки. Коллекторный переход, расположенный в высокоомном эпитаксиальном слое, имеет небольшую барьерную емкость и высокое пробивное напряжение. Транзисторы с такой структурой называют *эпитаксиально-планарными*. Они составляют основную часть транзисторов массового производства.

Методы повышения пробивного напряжения коллекторного перехода

Пробивное напряжение коллекторного (и эмиттерного) перехода планарного или эпитаксиально-планарного транзистора может оказаться низким, во-первых, из-за инородных включений и дефектов, которые могут быть в исходном кристалле полупроводника или появиться в области объемного заряда переходов в процессе их формирования. Особенно велика вероятность инородных включений, трещин и других дефектов вблизи поверхности кристалла. Они остаются там после различных обработок поверхности. Инородные включения, естественно, отличаются от

полупроводника значениями относительной диэлектрической проницаемости и удельного сопротивления. Именно поэтому инородные включения приводят к искажению картины электрического поля и уменьшению пробивного напряжения. Уменьшить вероятность появления инородных включений и дефектов можно только путем повышения качества исходных материалов и выполнения всех требований технологического процесса.

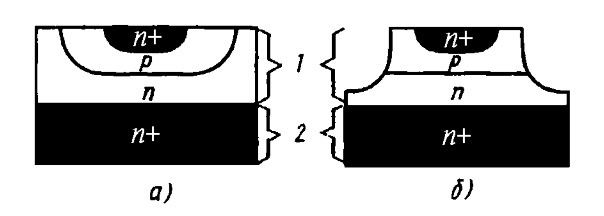


Рис. 4.49. Структуры эпитаксиально-планарного (*а*) и мезапланарного (*б*) транзисторов:

1 — эпитаксиальный слой кремния толщиной около 10 мкм; *2* — подложка из сильнолегированного кремния толщиной около 200 мкм

Во-вторых, пробивное напряжение коллекторного перехода может оказаться малым из-за поверхностного пробоя, обусловленного большой плотностью поверхностных состояний определенного типа и наличием в результате обогащенного слоя на поверхности высокоомной области, прилегающей к переходу (см. § 3.14). Эта разновидность пробоя перехода может существовать даже в кристаллах, поверхность которых не имеет механических нарушений.

Для уменьшения вероятности такого пробоя необходимо при обработке поверхности кристалла и при выращивании слоя диоксида кремния использовать методы, обеспечивающие наименьшую плотность поверхностных состояний.

В-третьих, пробивное напряжение перехода может быть понижено из-за кривизны *p-n*-перехода на его краях (рис. 4.49, *а*). Напряженность электрического поля всегда выше вблизи острых выступов, где сгущаются силовые линии тока. Глубина залегания *p-n*-переходов в транзисторах, изготовленных с использованием методов планарной технологии, обычно от десятых долей до единиц микрометров. Это означает, что радиус закругления *p-n*-переходов в местах их выхода на поверхность кристалла составляет в лучшем случае несколько микрометров. Уменьшение радиуса закругления краев *p-n*-перехода на порядок приводит к уменьшению пробивного напряжения также приблизительно на порядок.

Одним из методов устранения пробоя, связанного с кривизной *p-n*-перехода, — это селективное стравливание части кристалла, где расположены закругленные края перехода (рис. 4.49, *б*).

Структуру транзистора, полученную путем селективного травления, называют *мезаструктурой*, а транзисторы с мезаструктурой, *p-n*-переходы в которой сформированы методами планарной технологии, — *мезапланарными*. В структуре мезапланарного транзистора остается только плоская часть коллекторного перехода, имеющая значительно большее пробивное напряжение. При стравливании закругленных частей коллекторного перехода удаляются приповерхностные участки базы и коллектора, имеющие наибольшее количество нарушений кристаллической решетки и

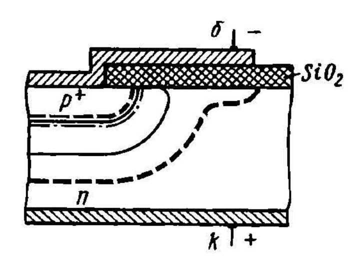


Рис. 4.50. Часть коллекторного перехода, выходящая на поверхность кристалла кремния, в транзисторе с расширенным базовым электродом (с металлизацией по оксиду):

——— — границы перехода без внешнего напряжения;

— — — — — границы перехода при обратном напряжении;

—·—·— — металлургическая граница перехода

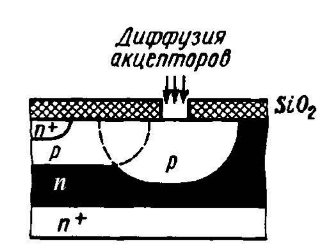


Рис. 4.51. Периферийная часть структуры эпитаксиально-планарного транзистора с диффузионным охранным кольцом, сформированным дополнительной диффузией акцепторов

других дефектов, что также способствует увеличению пробивного напряжения.

Другой метод повышения пробивного напряжения — создание расширенного металлического электрода к базовой области, простирающегося по слою диоксида кремния над коллекторным переходом и частично над коллекторной областью (рис. 4.50). При приложении к коллекторному переходу напряжения в обратном направлении под расширенной частью электрода базы в коллекторной области вблизи поверхности возникает обедненный основными носителями (в данном случае электронами) слой. Это приводит к большей толщине коллекторного перехода у поверхности кристалла по сравнению с толщиной плоской его части, а также к уменьшению кривизны перехода. Все это способствует увеличению пробивного напряжения коллекторного перехода вблизи поверхности кристалла.

Еще одним методом повышения пробивного напряжения коллекторного перехода является формирование охранного кольца в месте выхода перехода на поверхность кристалла (рис. 4.51). Для этого приходится проводить дополнительную диффузию

акцепторов (для транзистора *n-p-n*-типа) по периферии коллекторного перехода со всеми вспомогательными операциями окисления, фотолитографии и т. д., проводя эту диффузию на бóльшую глубину по сравнению с диффузией при формировании базовой области. В результате уменьшается кривизна коллекторного перехода в местах выхода его на поверхность кристалла или, другими словами, увеличивается радиус закругления в периферийной части перехода.

Наконец, повысить пробивное напряжение коллекторного перехода можно методом формирования диффузионных делительных колец (рис. 4.52). Диффузионные делительные кольца формируют одновременно с созданием базовой области транзистора на некотором расстоянии от нее. При обратном напряжении на коллекторном переходе происходит расширение этого перехода и его смыкание с переходом ближайшего диффузионного кольца. На этом первом диффузионном кольце устанавливается плавающий потенциал, а на *p-n*-переходе первого диффузионного кольца — обратное напряжение, которое по абсолютному значению меньше, чем на коллекторном переходе.

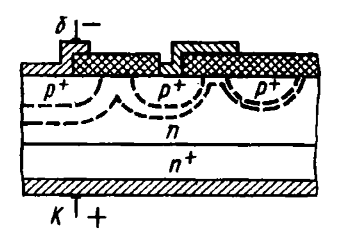


Рис. 4.52. Периферийная часть коллекторного перехода с двумя диффузионными делительными кольцами, созданными одновременно с формированием базовой области с электропроводностью *p*-типа

При наличии второго диффузионного кольца его *p-n*-переход смыкается с переходом первого диффузионного кольца. На первое диффузионное кольцо может быть нанесен электрод, расширенный по слою диоксида кремния до второго диффузионного кольца. Этот электрод, никуда не присоединенный и имеющий некоторый плавающий отрицательный потенциал по отношению к коллектору, будет способствовать смыканию переходов первого и второго диффузионных делительных колец.

Особенности конструкции и структуры мощных транзисторов

При разработке мощных транзисторов помимо тех задач, которые решаются при создании маломощных транзисторов, приходится решать следующие специфические проблемы:

1) мощные транзисторы работают при довольно больших токах, поэтому в них особенно сильно проявляются связанные с этим эффекты и это необходимо предусмотреть (см. § 4.5);

2) для обеспечения достаточной мощности в нагрузке обычно используют источники питания с большим напряжением, поэтому чаще всего мощные транзисторы должны быть рассчитаны на бóльшие пробивные напряжения, чем маломощные;

3) мощные транзисторы должны иметь большой коэффициент полезного действия и, в частности, малое падение напряжения на транзисторе в режиме насыщения, т. е. малое сопротивление насыщения;

4) конструкция мощного транзистора должна обеспечивать эффективный отвод рассеиваемой в нем теплоты;

5) значительный перегрев активных частей мощного транзистора при больших размерах применяемых в нем кристаллов полупроводника вызывает необходимость учета механических напряжений, которые могут возникать из-за различия темпера-

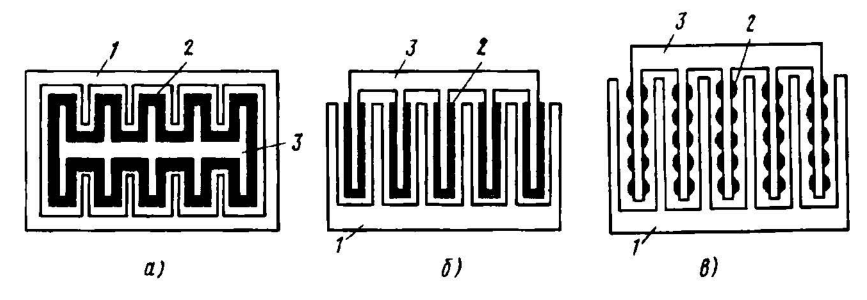


Рис. 4.53. Геометрия электродов мощного СВЧ-транзистора:
а — конструкция с гребенчатым эмиттером; *б* — многоэмиттерная конструкция с полосковыми эмиттерами; *в* — многоэмиттерная конструкция с круглыми эмиттерами; *1* — электрод базы; *2* — эмиттер; *3* — электрод эмиттера

турных коэффициентов линейного расширения полупроводника и других элементов конструкции.

Кроме того, мощные транзисторы должны быть и достаточно быстродействующими.

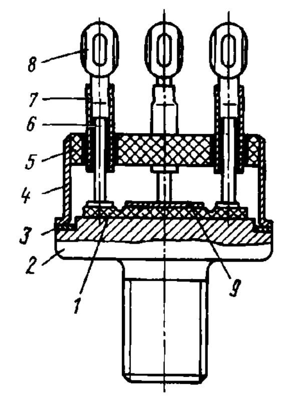
Для обеспечения больших токов в транзисторах необходимы большие площади эмиттера. Однако простое увеличение размеров эмиттера привело бы к значительному вытеснению тока на края эмиттерного перехода из-за падения напряжения на объемном сопротивлении базы. Планарная технология формирования транзисторных структур позволяет изготовлять эмиттеры сложной формы.

Одним из вариантов структуры мощных планарных транзисторов является гребенчатый эмиттер (рис. 4.53, *а*). Число зубьев «гребенки» может быть до нескольких десятков. Так как ширина каждого зубца небольшая, эффект вытеснения тока к краю эмиттера незначителен. Одновременно снижается и сопротивление базы, что увеличивает усиление по мощности на высокой частоте и снижает сопротивление насыщения. Однако при слишком узких и длинных эмиттерных полосках может сказаться падение напряжения вдоль них при больших токах.

Другими вариантами структур мощных транзисторов являются структуры с большим числом не связанных между собой эмиттеров в виде полосок (рис. 4.53, *б*) либо дисков (рис. 4.53, *в*).

Эти эмиттеры соединяют друг с другом путем металлизации поверх слоя диоксида кремния.

Для обеспечения лучшего теплоотвода кристалл полупроводника со структурой мощного транзистора припаивают к кристаллодержателю коллекторной стороной. Если же необходима изоляция коллектора от корпуса транзистора, то между кристаллом полупроводника и кристаллодержателем помещают изолирующую прокладку из бериллиевой керамики, обладающей хорошей теплопроводностью (рис. 4.54). Основание корпуса — кристаллодержатель — выполняют из меди. Так как теплоотвод от корпуса должен обычно осуществляться на шасси всего устройства или на радиатор, то для уменьшения теплового сопротивления нижнюю поверхность основания корпуса не окрашивают.



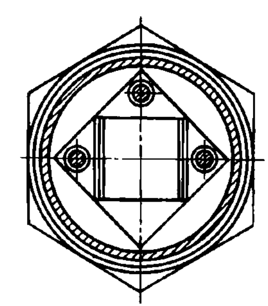


Рис. 4.54. Конструкция мощного СВЧ-транзистора:

1 — пластинка из бериллиевой керамики; 2 — фланец; 3 — кольцо; 4 — баллон; 5 — керамический изолятор; 6 — внутренний вывод; 7 — трубка; 8 — наружный вывод; 9 — кристалл полупроводника

Особенности конструкции и структуры СВЧ-транзисторов

Почти все СВЧ-транзисторы, как и остальные биполярные транзисторы, — это кремниевые эпитаксиально-планарные транзисторы. Низкоомная подложка исходной эпитаксиальной структуры обеспечивает малое сопротивление коллекторной области и ограничивает накопление носителей в этой области.

Для получения хороших частотных свойств должны быть сведены к минимуму значения паразитных параметров и размеры активных частей структуры транзистора. Поэтому толщина базы СВЧ-транзисторов составляет иногда менее 0,1 мкм, ширина эмиттерных полосок — менее 1 мкм. Для уменьшения объемного сопротивления базы при малой ее толщине необходима большая концентрация примесей в базе. Перспективным методом формирования базовых и эмиттерных областей малых размеров и с необходимой концентрацией примесей является метод ионной имплантации.

Формирование тонких базовых областей вызывает необходимость иметь эпитаксиальные слои кремния с практически совершенной кристаллической решеткой без дислокаций и других дефектов упаковки. Технологические операции, применяемые при создании транзисторных структур, не должны приводить к обра-

зованию таких дефектов. В противном случае будет велика вероятность закорачивания областей эмиттера и коллектора.

На работу транзистора в области сверхвысоких частот могут влиять емкости между выводами и индуктивность выводов. Поэтому применяют ленточные выводы с пониженной индуктивностью. Для уменьшеиия паразитных емкостей корпус СВЧ-траизистора вместе с радиатором, если он есть, желательно изолировать от коллекторной области, сохранив при этом хороший теплоотвод от коллекторного перехода. Достигают этого путем крепления кристалла кремиия с транзисторной структурой на пластинке из бериллиевой керамики (рис. 4.54).

Осиовные трудности возникают при создании мощных СВЧ-траизисторов, так как требоваиия к структуре высокочастотиого транзистора в осиовиом противоречат требованиям к структуре мощного траизистора. Так, одним из осиовных факторов, ограничивающих диапазон рабочих частот СВЧ-транзистора, является время пролета носителей заряда через коллекторный переход (см. § 4.9). Для повышения рабочей частоты толщину коллекторного перехода желательно уменьшить. Однако при этом оказывается малым и пробивное напряжение. В результате СВЧ-транзисторы с бо́льшими значениями граничной частоты имеют меньшее значение максимально допустимой мощности рассеяния. Лучшие из современных СВЧ-транзисторов с граничиыми частотами в иесколько гигагерц рассчитаны на максимально допустимую мощность рассеяния при непрерывной работе в несколько ватт.

§ 4.18. ОДНОПЕРЕХОДНЫЕ ТРАНЗИСТОРЫ

Одиопереходный транзистор — это полупроводниковый прибор с одним выпрямляющим электрическим переходом и тремя выводами, переключающие и усилительные свойства которого обусловлены модуляцией сопротивления базы в результате инжекции в нее неосновных иосителей заряда.

Структура однопереходного траизистора и его эквивалентиая схема показаны на рис. 4.55. Область эмиттера (область с электропроводностью p-типа) должна быть легирована сильнее, чем область базы (область с электропроводностью n-типа), для того чтобы при прямом включеиии эмиттерного перехода прямой ток через него имел в основном лишь дырочную составляющую. В этом случае из-за инжекции неосиовных носителей заряда в базу транзистора и из-за накопления основных иосителей, которые входят в базу через один из невыпрямляющих контактов к базе для компенсации инжектированного заряда неосновных носителей, будет происходить уменьшение сопротивлеиия базы (модуляция) и увеличение тока между иевыпрямляющими коитактами к базе или тока в цепи иагрузки.

Если на базовые выводы прибора подано межбазовое напряжение $U_{Б_2Б_1}$, то из-за прохождения тока $I_{Б_2}$ вдоль базы будет существовать продольное падение напряжения. Обозначим падение напряжения на части базы протяженностью l_1 (рис. 4.55, *а*) как $U_{вн}$. Это падение напряжения смещает *р-п*-переход в обратном направлении. Поэтому при напряжении на эмиттере $U_{ЭБ0} < U_{вн}$ *р-п*-переход смещен в обратном направлении и во входной цепи проходит небольшой обратный ток $I_{ЭБ0}$ (рис. 4.56, *а*).

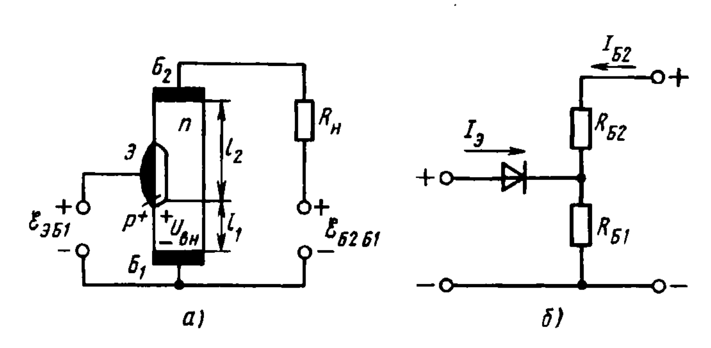


Рис. 4.55. Структура *(а)* и эквивалентная схема *(б)* однопереходного транзистора

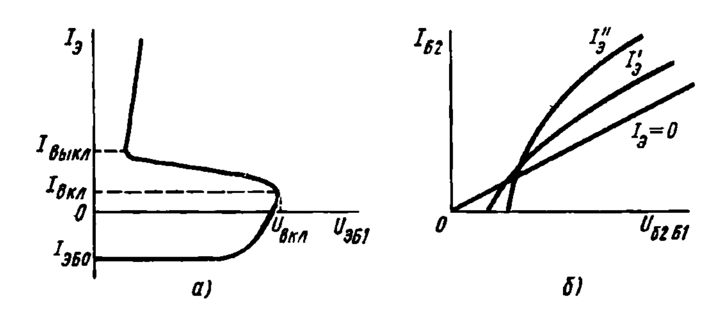


Рис. 4.56. Входная *(а)* и выходные *(б)* статические характеристики однопереходного транзистора

Если же входное напряжение, поданное на эмиттер относительно базы Б₁, превысит $U_{вн}$, то *р-п*-переход открывается и начинается инжекция неосновных носителей заряда (дырок) в базу. Вначале инжекция происходит только через часть *р-п*-перехода, расположенную ближе к первому контакту базы Б₁. В результате сопротивление части базы протяженностью l_1 уменьшается, что приводит к еще большему смещению *р-п*-перехода эмиттера в прямом направлении и к появлению на входной статической характери-

стике участка с отрицательным дифференциальным сопротивлением (рис. 4.56, *а*):

Таким образом, однопереходный транзистор может находиться в двух устойчивых состояниях: в закрытом, которое характеризуется относительно большими сопротивлениями между различными выводами однопереходного транзистора, и в открытом (или в состоянии насыщения), характеризующемся относительно малыми сопротивлениями. В открытом состоянии однопереходный транзистор будет находиться до тех пор, пока инжекция носителей заряда через эмиттерный переход будет поддерживать в базе избыточную концентрацию неосновных и основных носителей заряда, т. е. до тех пор, пока ток эмиттера будет превышать значение тока выключения $I_{выкл}$.

Выходные, или межбазовые, статические характеристики однопереходного транзистора представляют собой зависимости тока через второй базовый вывод от межбазового напряжения при определенных значениях тока эмиттера (рис. 4.56, *б*). При $I_Э=0$ выходная характеристика — прямая линия. При прямых токах эмиттера, отличных от нуля, выходные характеристики оказываются нелинейными, так как суммарное напряжение на эмиттерном переходе изменяется с изменением выходного тока $I_{Б_2}$.

Однопереходный транзистор, как всякий прибор с отрицательным дифференциальным сопротивлением, может быть использован в качестве переключателя, генератора и усилителя. При этом он может обеспечить усиление как по мощности и напряжению, так и по току.

Определим коэффициент передачи тока однопереходного транзистора. При большой напряженности электрического поля в базе можно пренебречь диффузией носителей заряда и учитывать только дрейфовые токи:

$$I_{Б2}=I_p+I_n=qS(p_{nб}\mu_p+n_{nб}\mu_n)E,$$

где S — площадь поперечного сечения базы (для тока $I_{Б2}$).

Допустим, что ток эмиттера увеличился на $\Delta I_Э$, тогда в базе увеличится концентрация дырок на $\Delta p_{nб}$ и одновременно для сохранения электрической нейтральности в базу из невыпрямляющих контактов войдет такое же количество основных носителей заряда $(\Delta n_{nб}=\Delta p_{nб})$.

Так как

$$\Delta I_Э=qS\Delta p_{nб}\mu_p E$$

и

$$\Delta I_{Б2}=qS(\Delta p_{nб}\mu_p+\Delta n_{nб}\mu_n)E,$$

то коэффициент передачи тока

$$k_i=\Delta I_{Б2}/\Delta I_Э=1+(\mu_n/\mu_p).$$

При этом предполагалось, что эффективность эмиттера близка к единице и коэффициент переноса носителей заряда по базе тоже

близок к единице в связи с малым временем дрейфа носителей в базе по сравнению с временем их жизни.

Так как подвижность электронов в большинстве полупроводников превышает подвижность дырок в том же материале, то для однопереходных транзисторов выгоднее применять исходный материал с электропроводностью n-типа.

Как видно, принцип действия однопереходного транзистора хотя и связан с инжекцией неосновных носителей в базу, но существенно отличается от принципа действия обычного биполярного транзистора. Кроме того, в связи с относительно большим объемом базы однопереходные тразисторы значительно уступают обычным биполярным транзисторам по частотным свойствам.

§ 4.19. НАДЕЖНОСТЬ ТРАНЗИСТОРОВ

Катастрофические отказы

Катастрофические отказы транзисторов являются чаще всего следствием либо недостаточной отработки их конструкции, либо нарушений в технологии. Как и в полупроводниковых диодах (см. § 3.32), катастрофические отказы транзисторов вызываются обрывами выводов, растрескиванием кристалла, пробоем переходов или оксидного слоя, короткими замыканиями.

Обрывы или перегорание внутренних выводов чаще всего случаются при использовании очень тонкой проволоки, которая бывает неоднородна по толщине. Такую проволоку применяют, например, для создания выводов планарных транзисторов. Вероятность такого отказа растет, если прибор подвергается воздействию вибрации, ударов и т. п., а также при изменениях температуры. Иногда наблюдаются обрывы вывода сплавного транзистора в месте припайки к навеске. Причиной такого обрыва может быть нарушение технологии пайки либо отсутствие термокомпенсирующего изгиба (тогда обрыв происходит при низкой температуре).

В планарных транзисторах короткие замыкания иногда возникают между металлическими слоями на поверхности оксида и полупроводником. Причиной таких замыканий могут быть отверстия в оксидном слое, появившиеся в результате нарушений слоя в процессе фотолитографии. Например, пылинка, осевшая на поверхность фоторезиста, приводит к тому, что под ней фоторезист не засвечивается, а следовательно, не полимеризуется. После травления оксида на месте, где находилась пылинка, получается отверстие. Не допустить дефекты такого рода можно, создав соответствующую чистоту в производственных помещениях.

Растрескивание кристалла чаще всего является следствием разницы в температурных коэффициентах линейного расширения полупроводника и электродного материала. Такой вид отказов характерен для мощных транзисторов, и, чтобы не допустить этих дефектов, применяют, например, термокомпенсирующие прокладки.

К особому виду катастрофических отказов транзисторов относятся те, которые связаны с эффектами шнурования тока (см. § 3.13). Так как в месте шнурования развивается очень высокая температура, она может приводить к плавлению электродного материала. Расплавленная область прорастает через базу и обусловливает короткое замыкание эмиттера с коллектором.

Условные отказы

Катастрофические отказы не являются основными для транзисторов — на их долю приходится примерно 20% всех наблюдаемых отказов. В основном же отказы транзисторов обусловлены постепенным изменением параметров сверх допустимых отклонений.

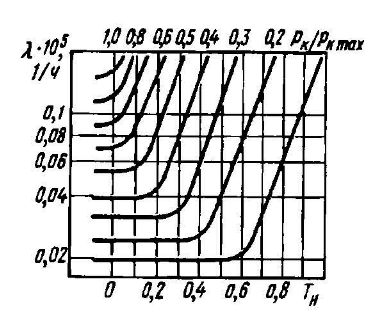


Рис. 4.57. Обобщенная зависимость интенсивности отказов транзисторов от нормализованной температуры при различных отношениях рабочей мощности рассеяния к максимально допустимой мощности на коллекторе

В транзисторах практически не наблюдается каких-либо объемных процессов, приводящих к изменению параметров. Токи транзистора являются чисто электронными, в них нет каких-либо ионных составляющих, а используемые примеси имеют при рабочей температуре транзистора ничтожные коэффициенты диффузии. Поэтому постепенные изменения параметров транзисторов связаны с явлениями на поверхности полупроводника.

В принципе влияние состояния поверхности на параметры транзистора подобно аналогичному влиянию у полупроводниковых диодов (см. § 3.14). Однако следует заметить, что помимо изменения токов переходов $I_{КБО}$, $I_{КЭО}$, $I_{КБК}$ для транзисторов характерно и изменение коэффициента передачи, связанное с изменением скорости поверхностной рекомбинации.

Одной из причин изменения состояния поверхности, вызывающей изменения параметров транзисторов, является наличие следов влаги в баллоне. Особенно сильно это сказывается на параметрах сплавных транзисторов. Стабильность параметров существенно улучшается, когда в корпусе помещается влагопоглотитель (цеолит), но даже такая мера иногда может быть недостаточно эффективной, если прибор подвергается циклическим изменениям температуры. При этом происходит перераспределение влаги между влагопоглотителем и поверхностью полупроводника, что достаточно для изменения параметров транзистора.

Другой причиной изменения состояния поверхности, характерной для планарных транзисторов, является движение ионов щелочных металлов в слое оксида, особенно заметное при подаче напряжения и высокой температуре.

К снижению надежности транзисторов могут приводить неблагоприятные условия работы — большая рассеиваемая мощность, повышенная температура, вибрация и т. д. Особенно сильным оказывается такое влияние, когда режим работы транзистора приближается к предельно допустимому. На рис. 4.57 показана обобщенная зависимость интенсивности отказов от отношения рабочей мощности рассеяния, отнесенной к максимально допустимой мощности на коллекторе, и от нормализованной температуры:

$$T_{н} = \frac{T_{раб} - T_{сн}}{T_{max} - T_{сн}},$$

где $T_{раб}$ — рабочая температура; T_{max} — максимально допустимая температура; $T_{сн}$ — температура, начиная с которой по техническим условиям требуется снижение электрического режима.

Таблица 4.2

Влияние внешних факторов на транзисторы

Внешнее воздействие	Вызываемые или ускоряемые процессы	Типичные дефекты
Повышенная температура	Высыхание защитных покрытий и их деформация Выделение газов Расплавление припоя Растрескивание кристалла Миграция ионов примесей и влаги Изменение параметров полупроводника	Необратимый тепловой пробой Увеличение обратных токов Потеря герметичности Обрывы и короткие замыкания внутренних выводов
Пониженная температура	Конденсация влаги Растрескивание кристалла Изменение электрофизических параметров полупроводника	Уменьшение пробивного напряжения Снижение коэффициентов передачи Потеря герметичности Обрывы и короткие замыкания внутренних выводов
Повышенная влажность	Адсорбция и абсорбция влаги Химические реакции с влагой Электролиз Коррозия	Изменение параметров Нестабильность параметров Коррозия выводов и корпуса Повреждения лакокрасочных покрытий
Резкие изменения температуры	Механические напряжения Растрескивание кристалла Растрескивание и деформация защитных покрытий	Потеря герметичности Обрывы и короткие замыкания внутренних выводов Изменение параметров
Пониженное давление	Ухудшение теплоотдачи	Перегрев кристалла
Механические ускорения	Механические напряжения Усталость	Обрывы и короткие замыкания выводов Потеря герметичности

Причиной отказов транзистора бывают механические перегрузки при постоянных ускорениях свыше 10 000g и ударных нагрузках свыше 1000g.

В табл. 4.2 обобщены факторы внешней среды, влияющие на параметры транзисторов.

Отказы полупроводниковых приборов могут происходить не только при их эксплуатации в неблагоприятных условиях, но и при хранении.

При изучении влияния радиации на надежность полупроводниковых приборов следует различать врéменные и необратимые изменения параметров.

Врéменные изменения наблюдаются во время воздействия γ-радиации и связаны в основном с появлением избыточной концентрации носителей заряда в объеме полупроводника, что обусловливает, например, рост обратного тока. Обратимым может быть также влияние ионизации газа в объеме баллона транзистора. Однако при этом возможно оседание ионов на поверхность полупроводника, что приводит к длительным изменениям параметров транзисторов.

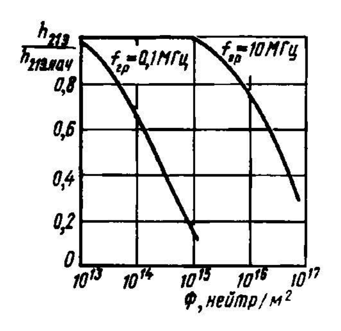


Рис. 4.58. Относительное изменение коэффициента передачи тока базы в зависимости от дозы нейтронов для низкочастотных и высокочастотных транзисторов

Облучение тяжелыми частицами (нейтронами) приводит к образованию различных дефектов в объеме полупроводника (см. § 9.10). При этом могут происходить необратимые изменения параметров транзисторов; например, из-за появления новых рекомбинационных ловушек может снижаться коэффициент передачи (рис. 4.58). Более высокая радиационная стойкость высокочастотных транзисторов связана с меньшей толщиной базы и, следовательно, с меньшей вероятностью появления в ней дефектов.

Контрольные вопросы

1. Каким образом в транзисторе происходит усиление электрических колебаний по мощности?
2. Почему транзистор, включенный по схеме с общим эмиттером, может обеспечить усиление по току?
3. По каким причинам в базе транзистора возникает электрическое поле?
4. Какие физические процессы происходят в транзисторе при его пробое, если он включен по схемам с общей базой и с общим эмиттером?
5. Как объяснить вид входных и выходных статических характеристик транзистора, включенного по схемам с общей базой и общим эмиттером?
6. Какие факторы определяют инерционность транзистора при его работе на высоких частотах?
7. Какие существуют системы малосигнальных параметров транзистора и в чем преимущества системы h-параметров?
8. Какие существуют эквивалентные схемы транзистора?
9. Что такое одномерная теоретическая модель транзистора?

10. В чем отличие коэффицента передачи тока эмиттера одномериой теоретической модели транзистора и коэффицента передачи тока эмиттера реального транзистора?

11. Как и почему коэффициент передачи тока эмиттера зависит от постоянного тока эмиттера?

12. Какие физические процессы происходят в транзисторе при его пробое, если он включен по схеме с общей базой и с общим эмиттером?

13. Почему при включении транзистора в схему, находящуюся под напряжением, надо первым подключить вывод базы, а потом остальные выводы?

14. Чем вызван вторичный пробой в транзисторах?

15. Какие явления могут возникать в транзисторе при больших плотностях тока эмиттера?

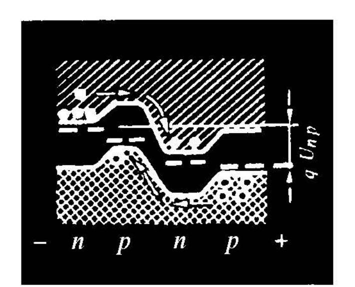
16. Какими параметрами можно характеризовать частотные свойства транзистора? Дайте определения этих параметров.

17. Почему при большом входном сигнале коллекторный переход транзистора может оказаться смещенным в прямом направлении, если полярность внешнего источника напряжения в цепи коллектора с сопротивлением нагрузки соответствует обратному включению этого перехода?

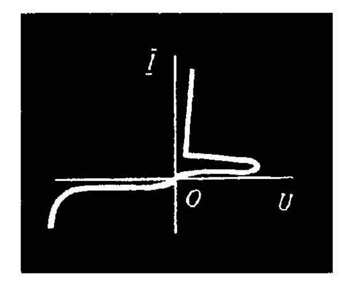
18. В чем особенности структуры и конструкции мощных транзисторов?

19. Какова природа шумов, возникающих в транзисторах?

20. Объясните принцип действия однопереходного транзистора.



Тиристоры



Тиристор — это полупроводниковый прибор с двумя устойчивыми состояниями, имеющий три (или более) выпрямляющих перехода, который может переключаться из закрытого состояния в открытое и наоборот.

§ 5.1. ДИОДНЫЕ ТИРИСТОРЫ

Диодный тиристор (динистор) — это тиристор, имеющий два вывода, через которые проходит как основной ток, так и ток управления.

Структура и принцип действия

Структура диодного тиристора состоит из четырех областей полупроводника с чередующимся типом электропроводности (рис. 5.1). Кроме трех выпрямляющих переходов диодный тиристор имеет два омических перехода. Один из омических переходов расположен между крайней *n*-областью и металлическим электродом, который называют *катодом*. Другой омический переход расположен между крайней *p*-областью и металлическим электродом, который называют *анодом*.

Вначале рассмотрим процессы, происходящие в тиристоре при подаче на него прямого напряжения, т. е. при положительном потенциале на аноде. В этом случае крайние *p-n*-переходы смещены в прямом направлении, поэтому их называют *эмиттерными*; средний *p-n*-переход смещен в обратном направлении, поэтому его называют *коллекторным*. Соответственно в таком приборе существуют две *эмиттерные области* (*n*- и *p*-эмиттеры) и две *базовые области* (*p*- и *n*-базы).

Бóльшая часть внешнего прямого напряжения падает на коллекторном переходе, так как он смещен в обратном направлении. Поэтому первый участок прямой ветви ВАХ тиристора похож на обратную ветвь ВАХ выпрямительного диода. С увеличением анодного напряжения, приложенного между анодом и катодом, увеличивается прямое напряжение и на эмиттерных переходах. Электроны, инжектированные из *n*-эмиттера в *p*-базу, диффундируют к коллекторному переходу, втягиваются полем коллекторного перехода и попадают в *n*-базу. Дальнейшему продвижению элек-

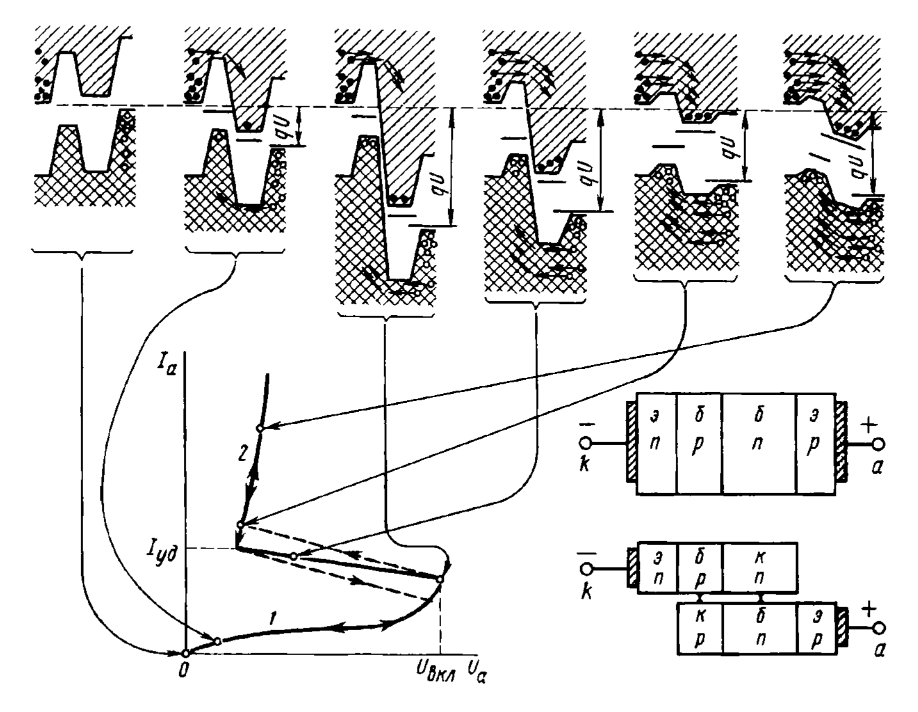


Рис. 5.1. Структура, ВАХ и энергетические диаграммы диодного тиристора

тронов по структуре тиристора препятствует небольшой потенциальный барьер правого эмиттерного перехода (рис. 5.1). Поэтому часть электронов, оказавшись в потенциальной яме *n*-базы, образует избыточный отрицательный заряд, который, понижая высоту потенциального барьера правого эмиттерного перехода, вызывает увеличение инжекции дырок из *p*-эмиттера в *n*-базу. Инжектированные дырки диффундируют к коллекторному переходу, втягиваются полем коллекторного перехода и попадают в *p*-базу. Дальнейшему их продвижению по структуре тиристора препятствует небольшой потенциальный барьер левого эмиттерного перехода. Следовательно, в *p*-базе происходит накопление избыточного положительного заряда, что обусловливает увеличение инжекции электронов из *n*-эмиттера. Таким образом, в структуре тиристора

существует положительная обратная связь по току — увеличение тока через один эмиттерный переход приводит к увеличению тока через другой эмиттерный переход.

Накопление неравновесных носителей в базовых областях равносильно дополнительной разности потенциалов на коллекторном переходе, которая в отличие от внешней разности потенциалов стремится сместить коллекторный переход в прямом направлении. Поэтому с увеличением тока через тиристор и, следовательно, с увеличением избыточных зарядов основных носителей заряда в базовых областях абсолютное значение суммарного напряжения на коллекторном переходе начнет уменьшаться. Ток через тиристор при этом будет ограничен только сопротивлением нагрузки и ЭДС источника питания. Высота потенциального барьера коллекторного перехода уменьшается до значения, соответствующего включению этого перехода в прямом направлении.

Таким образом, тиристор при подаче на него прямого напряжения может находиться в двух устойчивых состояниях: закрытом и открытом.

Закрытое состояние тиристора соответствует участку прямой ветви ВАХ между нулевой точкой и точкой переключения. Под *точкой переключения* понимают точку на ВАХ, в которой дифференциальное сопротивление равно нулю, а напряжение на тиристоре достигает максимального значения. В закрытом состоянии (участок *1* ВАХ на рис. 5.1) к тиристору может быть приложено большое напряжение, а ток при этом будет мал.

Открытое состояние тиристора соответствует низковольтному и низкоомному участку прямой ветви ВАХ. На рис. 5.1 открытому состоянию тиристора соответствует участок *2* ВАХ. Между первым и вторым участками ВАХ находится переходный участок, соответствующий неустойчивому состоянию тиристора. Особенно проявляется неустойчивость при относительно малом сопротивлении во внешней цепи тиристора. Тогда переключение тиристора из закрытого состояния в открытое и обратно происходит по штриховым линиям (рис. 5.1), наклон которых определяется обычно относительно малым сопротивлением нагрузки.

В открытом состоянии тиристор будет находиться до тех пор, пока за счет проходящего тока будет поддерживаться избыточный заряд в базах, необходимый для смещения коллекторного перехода в прямом направлении. Если же ток через тиристор уменьшить до некоторого значения, меньшего удерживающего тока $I_{уд}$, то в результате рекомбинации и рассасывания уменьшится количество неравновесных носителей заряда в базовых областях тиристора, коллекторный переход окажется смещенным в обратном направлении, произойдет перераспределение падений напряжения на выпрямляющих переходах тиристорной структуры, уменьшится инжекция из эмиттерных областей и тиристор перейдет в закрытое состояние (см. рис. 5.1). Таким образом, *удерживающий ток* тиристора — это минимальный ток, который необходим для поддержания тиристора в открытом состоянии.

Закрытое состояние

Структуру тиристора можно представить в виде двух транзисторов, соединенных между собой, как показано на рис. 5.1. Постоянный ток коллектора этих транзисторов можно выразить через эмиттерные токи по аналогии с (4.47), пользуясь параметрами одномерной теоретической модели транзистора:

$$I_{п2} = \bar{\alpha}_1 I_{п1} + \bar{\alpha}_2 I_{п3} + I_{КБО}, \qquad (5.1)$$

где $I_{п1}$, $I_{п2}$ и $I_{п3}$ — токи через первый, второй и третий *p-n*-переходы; $\bar{\alpha}_1$ и $\bar{\alpha}_2$ — статические коэффициенты передачи токов эмиттера одномерных теоретических моделей первого и второго транзисторов; $I_{КБО}$ — обратный ток коллектора, который является общим для обоих транзисторов, составляющих структуру тиристора.

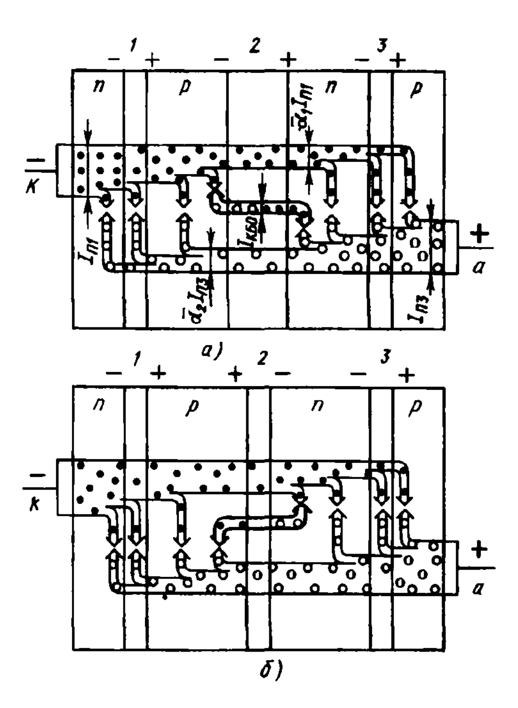


Рис. 5.2. Схематическое изображение потоков носителей заряда через структуру диодного тиристора в закрытом (*а*) и открытом (*б*) состояниях

Распределение потоков электронов и дырок в тиристорной структуре, находящейся в закрытом состоянии, схематично показано на рис. 5.2, *а*. На этом рисунке кроме основных потоков носителей заряда из эмиттерных областей в прилегающие базовые области учтена инжекция носителей из базовых областей в эмиттерные и рекомбинация носителей в эмиттерных переходах. В закрытом состоянии тиристора из всего потока инжектированных в каждую базу неосновных носителей заряда только меньшая часть доходит до коллекторного перехода. Основной механизм образования обратного тока коллектора — генерация носителей заряда в коллекторном переходе. На рисунке показана полярность напряжений, падающих на каждом из трех *p-n*-переходов от внешнего источника питания.

Для двухэлектродной структуры диодного тиристора из-за необходимости выполнения баланса токов полные токи через все переходы должны быть равны между собой:

$$I_{п1} = I_{п2} = I_{п3} = I_{а}. \qquad (5.2)$$

Тогда анодный ток тиристора

$$I_а = I_{КБО}/(1-\bar{\alpha}), \qquad (5.3)$$

где $\bar{\alpha} = \bar{\alpha}_1 + \bar{\alpha}_2$ — суммарный статический коэффициент передачи тока тиристорной структуры.

Выражение (5.3) представляет собой уравнение ВАХ диодного тиристора в закрытом состоянии. Напомним, что статический коэффициент передачи тока эмиттера транзистора растет с увеличением тока эмиттера в результате уменьшения влияния рекомбинации в эмиттерном переходе и появления электрического поля в базе из-за увеличения градиента концентрации носителей заряда. Коэффициент передачи тока эмиттера растет также с увеличением напряжения на коллекторном переходе в результате уменьшения толщины базы и увеличения коэффициента лавинного размножения в коллекторном переходе. Эти четыре физических фактора вызывают рост суммарного статического коэффициента передачи тока тиристорной структуры при увеличении напряжения и соответственно тока в закрытом состоянии тиристора.

При достижении суммарным статическим коэффициентом передачи значения, равного единнце, в соответствии с (5.3) анодный ток через тиристор устремляется в бесконечность, т. е. происходит переключение диодного тиристора из закрытого состояния в открытое. Во время переключения ток через тиристор, конечно, должен быть ограничен сопротивлением нагрузки, иначе тиристор может выйти из строя.

Условие переключения тиристора

В точке переключения тиристора дифференциальное сопротивление равно нулю. Выясним условие, при котором дифференциальное сопротивление тиристора может стать равным нулю. До переключения тиристора в открытое состояние практически все напряжение, приложенное к тиристору, падает на коллекторном переходе. Дифференцируя (5.1) по напряжению с учетом (5.2) и считая

$$\frac{d\bar{\alpha}(I,\ U)}{dU} = \frac{\partial\bar{\alpha}}{\partial I}\frac{\partial I}{\partial U} + \frac{\partial\bar{\alpha}}{\partial U},$$

получим

$$r = \frac{dU_а}{dI_а} = \frac{1-\left(\bar{\alpha}_1 + I_а\dfrac{\partial\bar{\alpha}_1}{\partial I_а}\right)-\left(\bar{\alpha}_2 + I_а\dfrac{\partial\bar{\alpha}_2}{\partial I_а}\right)}{\dfrac{\partial I_{КБО}}{\partial U_а} + I_а\left(\dfrac{\partial\bar{\alpha}_1}{\partial U_а} + \dfrac{\partial\bar{\alpha}_2}{\partial U_а}\right)}. \qquad (5.4)$$

В скобках числителя (5.4) стоят выражения для дифференциального коэффициента передачи тока эмиттера одномерных теоретических моделей транзисторов. Действительно, $I_К = \bar{\alpha}I_Э + I_{КБО}$. Отсюда дифференциальный коэффициент передачи тока эмиттера одномерной теоретической модели транзистора $\alpha =$

$= \partial I_К / \partial I_Э = \bar{\alpha} + I_Э(\partial \bar{\alpha} / \partial I_Э)$. Таким образом, из (5.4) следует, что переключение тиристора из закрытого состояния в открытое должно произойти при условии равенства единице суммарного дифференциального коэффициента передачи тока тиристорной структуры, т. е.

$$\alpha = \alpha_1 + \alpha_2 = 1. \qquad (5.5)$$

Обычно это условие удовлетворяется раньше, чем условие равенства единице статических коэффициентов передачи тока первого и второго транзисторов, составляющих тиристорную структуру, так как дифференциальные коэффициенты передачи несколько больше статических.

Ничтожное превышение дифференциального коэффициента передачи тока над единицей означает, что приращение тока коллектора больше, чем приращение тока эмиттера. Именно при таком условии *p*-база заряжается положительно, а *n*-база — отрицательно (об этом шла речь в начале этого параграфа). Избыточные заряды в базовых областях уменьшают напряжение на коллекторном переходе, а следовательно, и на всей тиристорной структуре, что соответствует переходному участку ВАХ тиристора — участку отрицательного дифференциального сопротивления.

При переключении диодного тиристора из закрытого состояния в открытое из-за роста проходящего тока суммарный дифференциальный коэффициент передачи тока увеличивается. Одновременно уменьшение напряжения на коллекторном переходе вызывает уменьшение того же коэффициента. Поэтому соотношение (5.5) можно считать не только условием переключения тиристора из закрытого состояния в открытое, но и уравнением ВАХ на переходном ее участке. В действительности сумма дифференциальных коэффициентов передачи тока транзисторных структур, составляющих тиристорную структуру, во время переключения несколько превышает единицу.

Открытое состояние

В открытом состоянии значение суммарного коэффициента передачи тока тиристорной структуры превышает единицу, т. е. бóльшая часть носителей заряда, инжектированных из эмиттерных областей, доходит до коллекторного перехода. Для открытого состояния при установившемся токе через диодный тиристор также должен сохраниться баланс токов. Поэтому необходимо предположить инжекцию дырок через коллекторный переход из *p*-базы в *n*-базу и инжекцию электронов в другом направлении (рис. 5.2, *б*). Это предположение, соответствующее смещению коллекторного перехода в прямом направлении, позволяет понять существование равенства полных потоков носителей заряда разных знаков во всех сечениях тиристорной структуры при установившемся режиме в открытом состоянии.

На рис. 5.2, *б* показана поляриость напряжеиий иа всех трех *p-n*-переходах тиристора в открытом состояиии. Коллекторный переход смещеи в прямом направлении из-за избыточиых зарядов осиовных иосителей в базовых областях, накоплениых там в процессе переключеиия тиристора.

Падение напряжеиия иа диодиом тиристоре в открытом состояиии представляет собой сумму напряжений иа всех *p-n*-переходах (с учетом ииверсии поляриости напряжения иа коллекториом переходе), иа объемиом сопротивлеиии различиых областей (в осиовиом слаболегированиой базы) и иа омических переходах.

Обратное иапряжение на диодиом тиристоре

При обратном иапряжении иа тиристоре, т. е. при отрицательном потеициале иа аиоде, эмиттерные переходы смещены в обратиом иаправлеиии, а коллекториый переход — в прямом. В этом случае нет условий для переключеиия тиристора, а обратиое иапряжеиие может быть ограиичеио либо лавииным пробоем эмиттерных переходов, либо эффектом смыкания переходов в результате расширеиия одного из эмиттерных переходов, смещениых в обратиом направлении, на всю толщииу слаболегированиой базы.

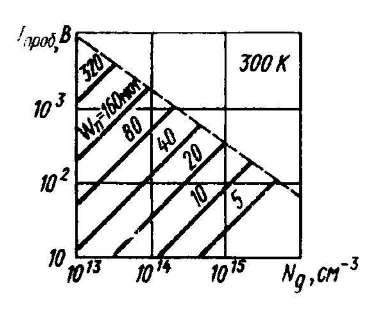


Рис. 5.3. Зависимости пробивного напряжения при лавинном пробое кремниевого несимметричного резкого *p-n*-перехода (штрихован линия) и при эффекте смыкаиия переходов тиристора (сплошные линии) от концентрации доноров в слаболегированной *n*-базе для разных толщин этой базы

На рис. 5.3 показаны расчетные значения пробивиого напряжения при лавинном пробое и при эффекте смыкания в зависимости от концентрации доиоров в слаболегированиой *n*-базе для разиых толщин этой базы. Для каждой толщииы базы обратиое напряжеиие иа тиристоре ограничено лавииным пробоем при больших коицеитрациях примеси в базе, так как в этом случае будет мала толщина эмиттериого перехода (см. § 3.11 и, в частиости, рис. 3.17, *а*). В тонком переходе иапряженность электрического поля, иеобходимая для ударной иоиизации, возиикиет при меньших иапряжениях. При малых коицеитрациях примеси в базе обратиое напряжение иа тиристоре будет ограиичено эффектом смыкаиия переходов, так как толщииа *p-n*-перехода тем больше, чем меиьше концентрация примеси в базе [см. соотиошение (2.27) в § 2.6].

Отметим, что рассмотрениые физические явлеиия ограиичивают не только обратиые иапряжеиия, ио и напряжеиия включеиия тиристоров. Только при прямом напряжении иа тиристоре ударная иоиизация может происходить в коллекторном переходе,

который смещен в обратном направлении при закрытом состоянии тиристора, а эффект смыкания может происходить из-за расширения коллекторного перехода.

§ 5.2. ДИОДНЫЙ ТИРИСТОР С ЗАШУНТИРОВАННЫМ ЭМИТТЕРНЫМ ПЕРЕХОДОМ

Переключение тиристора из закрытого состояния в открытое в соответствии с (5.5) происходит при возрастании суммарного дифференциального коэффициента передачи тока до единицы. В то же время в каждой из транзисторных структур, составляющих тиристор, коэффициенты передачи тока эмиттера могут быть близки к единице уже при малых напряжениях и токах. Для уменьшения начального значения коэффициента передачи одну из базовых областей всех тиристоров делают относительно толстой (до 200 мкм).

Для уменьшения коэффициента передачи тока эмиттера другого транзистора его эмиттерный переход шунтируют объемным сопротивлением прилегающей базовой области (рис. 5.4). Такое шунтирование осуществляют путем нанесения одного из основных электродов (например, катода) не только на эмиттерную область, но и частично на поверхность прилегающей базовой области. Шунтирование обеспечивает малые значения коэффициента передачи тока при малых напряжениях на тиристоре, так как почти весь ток при этом проходит по шунтирующему сопротивлению базы, минуя левый (рис. 5.4) эмиттерный переход в связи с его относительно большим сопротивлением при малых напряжениях. При больших напряжениях на тиристоре сопротивление левого эмиттерного перехода становится меньше шунтирующего сопротивления базы. Это значит, что теперь почти весь ток будет проходить через эмиттерный переход и будет вызван инжекцией неосновных носителей заряда в прилегающую базовую область.

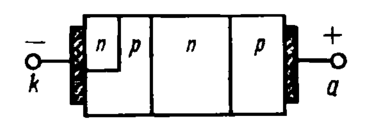


Рис. 5.4. Структура диодного тиристора с зашунтированным эмиттерным переходом

Шунтирование, во-первых, дает возможность создавать тиристоры с большими значениями напряжения включения. Во-вторых, при шунтировании эмиттерного перехода получается более резкая зависимость коэффициента передачи тока от напряжения и от тока. Поэтому тиристор с зашунтированным эмиттерным переходом будет иметь так называемую жесткую характеристику переключения, т. е. будет переходить из закрытого в открытое состояние каждый раз при одном и том же напряжении включения. Наоборот, при слабой зависимости коэффициента передачи тока от напряжения и от тока переключение тиристора из закрытого состояния в открытое может происходить при различных

значениях напряжения включения, т. е. тиристор в этом случае будет иметь так называемую мягкую характеристику переключения.

§ 5.3. ТРИОДНЫЕ ТИРИСТОРЫ

Триодный тиристор (тринистор) — это тиристор, имеющий два основных и один управляющий вывод.

Для переключения триодного тиристора из закрытого состояния в открытое также необходимо накопление неравновесных носителей заряда в базовых областях. В диодном тиристоре при увеличении напряжения на нем до напряжения включения это накопление неравновесных носителей заряда происходит обычно либо из-за увеличения уровня инжекции через эмиттерные переходы, либо из-за ударной ионизации в коллекторном переходе. В триодном тиристоре, имеющем управляющий вывод от

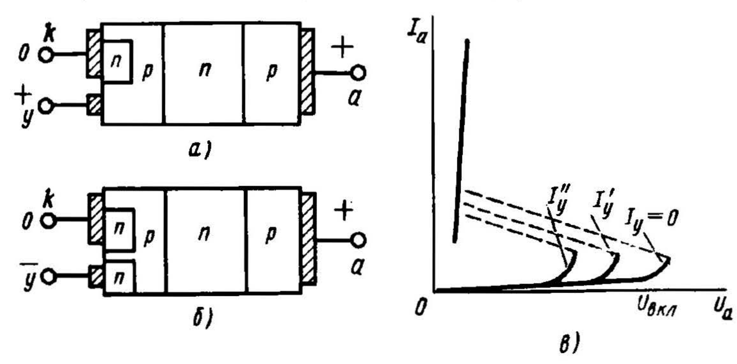


Рис. 5.5. Схематическое изображение структур триодных тиристоров с омическим переходом между управляющим электродом и базой *(а)*, с дополнительным *p-n*-переходом под управляющим электродом *(б)* и ВАХ триодного тиристора при различных токах ($I_y''>I_y'>I_y$) через управляющий электрод *(в)*

одной из базовых областей с омическим переходом между управляющим электродом и базой (рис. 5.5, *а*), уровень инжекции через прилегающий к этой базе эмиттерный переход можно увеличить путем подачи положительного по отношению к катоду напряжения на управляющий электрод. Поэтому триодный тиристор можно переключить из закрытого состояния в открытое в необходимый момент времени даже при небольшом анодном напряжении (рис. 5.5, *в*).

Переключение триодного тиристора с помощью подачи прямого напряжения на управляющий электрод или тока через этот электрод можно представить с другой точки зрения как перевод транзисторной *n-p-n*-структуры в режим насыщения при большом токе базы. При этом коллекторный переход транзисторной структуры (он же и коллекторный переход тиристора) смещается в прямом направлении.

Баланс токов в триодном тиристоре можно записать по аналогии с (5.1), но с учетом того, что через левый эмиттерный переход (рис. 5.5, *а*) проходит сумма токов основного и управляющего:

$$I_а = \bar{\alpha}_1(I_а + I_у) + \bar{\alpha}_2 I_а + I_{КБО} \qquad (5.6)$$

или

$$I_а(1 - \bar{\alpha}_1 - \bar{\alpha}_2) = I_{КБО} + \bar{\alpha}_1 I_у. \qquad (5.7)$$

Таким образом, уравнение ВАХ триодного тиристора в закрытом состоянии:

$$I_а = \frac{I_{КБО}}{1 - \bar{\alpha}_1 - \bar{\alpha}_2} + \frac{\bar{\alpha}_1}{1 - \bar{\alpha}_1 - \bar{\alpha}_2} I_у, \qquad (5.8)$$

где $\bar{\alpha}_1 + \bar{\alpha}_2 < 1$, а анодный ток $I_а$ зависит от управляющего тока $I_у$ (рис. 5.5, *в*).

Условие переключения триодного тиристора из закрытого состояния в открытое можно получить аналогично условию переключения диодного тиристора (см. § 5.1), т. е. после дифференцирования и преобразований получим

$$\alpha_1 + \alpha_2 + I_у \frac{\partial \bar{\alpha}_1}{\partial I_а} = 1. \qquad (5.9)$$

В условии переключения триодного тиристора (5.9) дифференциальный коэффициент передачи тока эмиттера одномерной теоретической модели первой транзисторной структуры α_1 зависит от напряжения на коллекторном переходе, а также от основного и управляющего токов. Аналогичный коэффициент второй транзисторной структуры α_2 зависит только от напряжения на коллекторе и от основного тока.

Из формулы (5.9) видно, что напряжение включения триодного тиристора зависит от управляющего тока. Формула может быть справедлива при меньших напряжениях на аноде тиристора, если через управляющий электрод будут проходить большие значения управляющего тока в прямом направлении. Кроме того, из условия (5.9) можно сделать вывод о целесообразности осуществления управляющего вывода от тонкой базы триодного тиристора, так как управлять коэффициентом передачи тока эмиттера транзисторной структуры с тонкой базой значительно легче, чем с толстой базой.

В открытом состоянии через тиристор проходит большой анодный ток, поэтому управляющий ток практически не оказывает влияния на участок ВАХ, соответствующий открытому состоянию триодного тиристора.

Управляющий электрод может быть сделан не только с омическим переходом между электродом и базовой областью, но и с дополнительным *p-n*-переходом (рис. 5.5, *б*). При определенной полярности напряжения на управляющем электроде относительно катода дополнительный переход окажется смещенным в пря-

мом направлении, через него будет происходить инжекция неосновных носителей заряда (для прилегающей базы) с последующим накоплением в другой базовой области. Такой процесс может привести к переключению триодного тиристора в открытое состояние.

Тиристор, у которого управляющий электрод соединен с *n*-областью, ближайшей к катоду, и который переводится в открытое состояние при подаче на управляющий электрод отрицательного по отношению к катоду сигнала, называют тиристором с инжектирующим управляющим электродом *n*-типа.

Триодный тиристор, структура которого представлена на рис. 5.5, *б*, можно рассматривать также как два диодных тиристора, имеющих общие анод, одну эмиттерную и обе базовые области. Структура основного тиристора выполнена с зашунтированным эмиттерным переходом. Поэтому напряжение включения основного тиристора больше, чем управляющего. При подаче отрицательного потенциала на управляющий электрод, т. е. на катод управляющего тиристора, можно переключить его из закрытого состояния в открытое. А так как обе структуры имеют общие области, то при переключении управляющего тиристора основной тиристор также окажется в открытом состоянии.

§ 5.4. ТИРИСТОРЫ, ПРОВОДЯЩИЕ В ОБРАТНОМ НАПРАВЛЕНИИ

Тиристор, проводящий в обратном направлении, — это тиристор, который при отрицательном анодном напряжении оказывается открытым в обратном направлении.

В предыдущих параграфах этой главы были рассмотрены диодные и триодные тиристоры, не проводящие в обратном направлении. Обратные токи таких тиристоров были малы вплоть до пробивного напряжения.

Тиристоры, проводящие в обратном направлении, могут быть диодными и триодными. Общей особенностью их структуры является шунтирование всех эмиттерных переходов объемными сопротивлениями прилегающих базовых областей (рис. 5.6, *а, б*). Для уменьшения шунтирующего сопротивления высокоомной базы (*n*-базы на рис. 5.6) ее поверхностный слой, прилегающий к эмиттерному переходу, дополнительно легируют соответствующей примесью.

В результате такого шунтирования при обратном напряжении на тиристоре (отрицательный потенциал на аноде) все эмиттерные переходы оказываются закороченными относительно малыми сопротивлениями, а коллекторный переход — смещенным в прямом направлении. Поэтому обратные токи через тиристоры, проводящие в обратном направлении, велики при малых обратных напряжениях (рис. 5.6, *в*).

Следствием активного шунтирования всех эмиттерных переходов является также то, что оба коэффициента передачи тока эмиттера транзисторных структур, составляющих тиристор, оказываются малыми. Поэтому при прямом напряжении на тиристоре основным физическим процессом, приводящим к накоплению неравновесных носителей заряда в базовых областях и к переключению тиристора из закрытого состояния в открытое, будет ударная ионизация в коллекторном переходе. При расчете и анализе характеристик и параметров таких тиристоров (да и

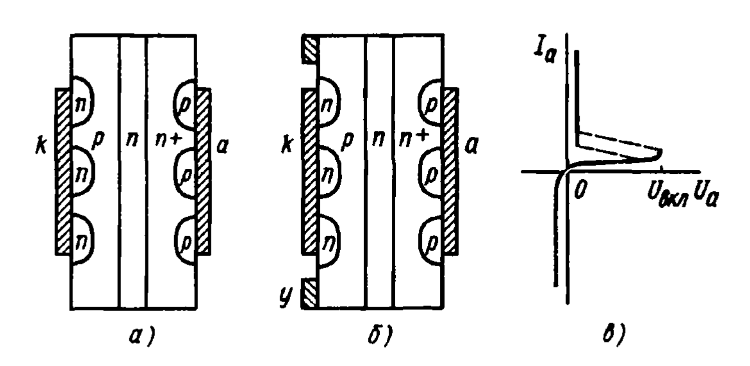


Рис. 5.6. Структуры диодного (*а*) и триодного (*б*) тиристоров, проводящих в обратном направлении, с распределенным шунтированием эмиттерных переходов и ВАХ тиристора, проводящего в обратном направлении (*в*)

многих других тиристоров, не проводящих в обратном направлении) надо учитывать в основном изменение коэффициента лавинного размножения M с напряжением на коллекторном переходе [см. (3.70)]. Например, условие переключения диодного тиристора из закрытого состояния в открытое (5.5) можно раскрыть следующим образом:

$$\alpha = \alpha_0 M = \frac{\alpha_0}{1-(U_{вкл}/U_{проб})^b} = 1,$$

где α_0 — суммарный дифференциальный коэффициент передачи тока тиристорной структуры при малых напряжениях, т. е. без учета лавинного размножения в коллекторном переходе.

Отсюда напряжение включения диодного тиристора

$$U_{вкл} = U_{проб} \sqrt[b]{1-\alpha_0}. \tag{5.10}$$

Преобладание лавинного размножения носителей заряда в коллекторном переходе над другими физическими процессами, которые могут приводить к переключению тиристора из закрытого состояния в открытое, обусловливает ряд особенностей в свойствах, отмеченных в последних параграфах этой главы.

Симметричный диодный тиристор (диак) — это диодный тиристор, способный переключаться как в прямом, так и в обратном направлениях.

Симметричный триодный тиристор (триак) — это триодный тиристор, который при подаче сигнала на его управляющий электрод включается как в прямом, так и в обратном направлениях.

Структура симметричного диодного тиристора состоит из пяти областей с чередующимся типом электропроводности, которые образуют четыре *p-n*-перехода (рис. 5.7, *а*). Крайние переходы зашунтированы объемными сопротивлениями прилегающих областей с электропроводностью *p*-типа.

Если на такой тиристор подать напряжение положительным потенциалом на область n_1 и отрицательным потенциалом на область n_3, то *p-n*-переход *1* окажется смещенным в обратном

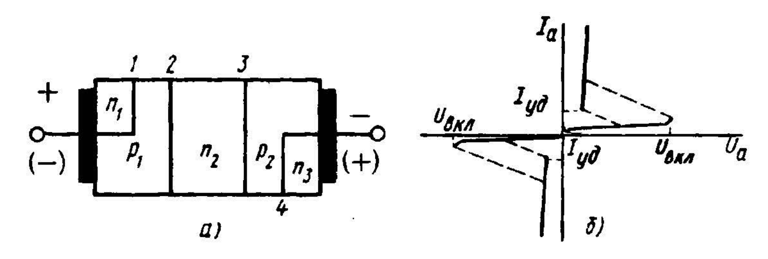


Рис. 5.7. Структура симметричного тиристора (*а*) и его ВАХ (*б*)

направлении и ток, проходящий через него, будет пренебрежимо мал. Весь ток через тиристор при такой полярности приложенного напряжения будет проходить по шунтирующему сопротивлению области p_1. Четвертый *p-n*-переход будет смещен в прямом направлении, и через него будет происходить инжекция электронов. При выбранной полярности внешнего напряжения рабочая часть тиристора представляет собой структуру *p-n-p-n*, в которой могут происходить те же процессы, что и в обычном диодном тиристоре, приводящие к переключению его из закрытого состояния в открытое и обратно.

При перемене полярности внешнего напряжения четвертый *p-n*-переход окажется смещенным в обратном направлении и, обладая поэтому большим сопротивлением, будет зашунтирован относительно малым сопротивлением области p_2. Следовательно, при такой полярности внешнего напряжения рабочая часть тиристора представляет собой структуру *n-p-n-p*, способную переключаться из закрытого состояния в открытое и обратно.

Таким образом, симметричный диодный тиристор можно представить в виде двух диодных тиристоров, включенных встречно и шунтирующих друг друга при разных полярностях приложенного напряжения. Вольт-амперная характеристика тако-

го тиристора получается одинаковой при разных полярностях приложенного напряжения (рис. 5.7, *б*).

Симметричные триодные тиристоры могут иметь структуру, способную переключаться из закрытого в открытое состояние либо при токе управляющего электрода определенного направления, либо при токе управляющего электрода любого направления (рис. 5.8). В последнем случае не только основные электроды должны обеспечить шунтирование прилегающих к ним крайних *p-n*-переходов, но и управляющий электрод должен иметь омический переход как с *p*-областью, так и с дополнительной *n*-областью. При этих условиях подача различных по знаку потенциалов на управляющий электрод по отношению к расположенному вблизи основному электроду будет либо изменять потенциал *p*-области, либо обеспечивать инжекцию электронов из дополнительной *n*-области.

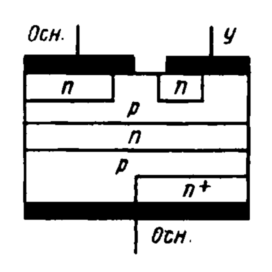


Рис. 5.8. Структура симметричного тиристора, переключаемого из закрытого состояния в открытое током управляющего электрода любого направления

§ 5.6. СПОСОБЫ УПРАВЛЕНИЯ ТИРИСТОРАМИ

Включение тиристоров

Включение тиристора путем медленного увеличения напряжения между основными электродами до напряжения включения $U_{вкл}$. Включение тиристора — это его переключение из закрытого состояния в открытое. Медленное увеличение напряжения между основными электродами тиристора до напряжения включения предполагалось при рассмотрении физических причин и условий включения различных диодных тиристоров в предыдущих параграфах. Таким способом можно, конечно, включить и триодный тиристор.

Включение тиристора с помощью тока управления. Как было показано в § 5.3, увеличение тока через один из эмиттерных переходов из-за подачи соответствующего напряжения на управляющий электрод приводит к накоплению неравновесных носителей заряда в базовых областях тиристора и к включению его при напряжении между основными электродами, значительно меньшем, чем напряжение включения при разомкнутой цепи управляющего электрода. Процесс накопления неравновесных носителей заряда в базовых областях происходит не мгновенно, поэтому для включения тиристора необходимо, чтобы импульс управляющего тока имел определенную длительность и амплитуду.

Время включения по управляющему электроду тиристора $t_{у.вкл}$ — это интервал времени между моментом в начале отпирающего импульса управляющего электрода, соответствующим 0,1 его амплитуды, и моментом, когда основное напряжение падает до 0,1 значения разности напряжений в закрытом и открытом состояниях тиристора или когда основной ток увеличится до 0,9 значения тока в открытом состоянии (рис. 5.9).

Время включения по управляющему электроду тиристора можно представить в виде суммы времени задержки по управляющему электроду и времени нарастания для тиристора.

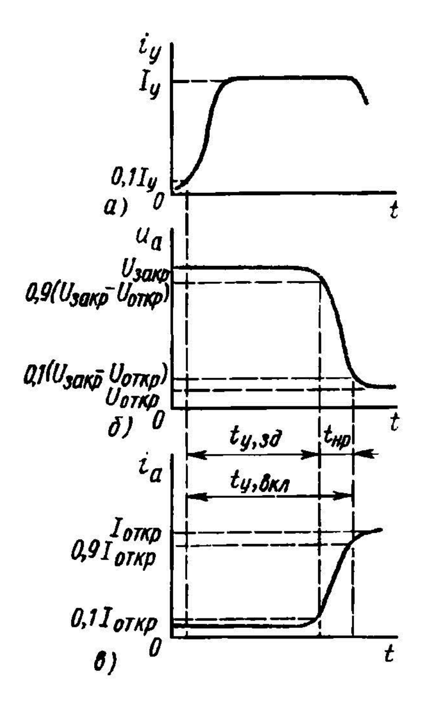


Рис. 5.9. Временные зависимости тока управляющего электрода (*а*), основного напряжения на тиристоре (*б*) и основного тока через тиристор (*в*), характеризующие процесс его включения

Время задержки по управляющему электроду тиристора $t_{у.зд}$ — это интервал времени между моментом в начале отпирающего импульса управляющего электрода, соответствующим 0,1 его амплитуды, и моментом, когда основное напряжение падает до 0,9 значения разности напряжений в закрытом и открытом состояниях тиристора или когда основной ток увеличивается до 0,1 его значения в открытом состоянии. ***Время нарастания*** для тиристора $t_{нр}$ — это интервал времени, в течение которого основной ток увеличивается от 0,1 до 0,9 значения тока в открытом состоянии или основное напряжение падает от 0,9 до 0,1 значения разности напряжений в закрытом и открытом состояниях тиристора (рис. 5.9).

Несмотря на условность определения всех перечисленных параметров переходного процесса включения тиристора, можно считать, что время задержки по управляющему электроду тиристора определяется временем перезаряда барьерной емкости эмиттерного перехода, а также временем прохождения инжектированных носителей заряда через базовую область и коллекторный переход. Время нарастания для тиристора определяется инерционностью процесса накопления неравновесных носителей заряда в базовых областях и инерционностью перезаряда барьерной емкости коллекторного перехода.

Интервал времени, определяемый временем нарастания для тиристора, изменением основного тока, проходящего через тиристор, и напряжения между основными электродами, соответ-

ствует переходному участку ВАХ. Для него сумма дифференциальных коэффициентов передачи тока эмиттера транзисторных структур, составляющих тиристор, должна быть равна единице. Обычно тиристор включен в цепь, имеющую сопротивление, меньшее абсолютного значения отрицательного дифференциального сопротивления тиристора на переходном участке его ВАХ. Поэтому в период времени нарастания при тех же напряжениях токи, проходящие через тиристор в реальной схеме, превышают значения токов переходного участка ВАХ тиристора. При этом суммарный дифференциальный коэффициент передачи тока тиристорной структуры превышает единицу, что соответствует активному этапу включения тиристора. Если в это время отключить управляющий электрод, то тиристор «самостоятельно» перейдет в открытое состояние. Значит, длительность импульса тока управляющего электрода, необходимая для включения тиристора, должна быть больше времени задержки.

Завершением процесса включения тиристора считают момент изменения полярности напряжения на коллекторном переходе.

Необходимо отметить, что во время переходных процессов через тиристор проходят большие токи при больших напряжениях на нем, что приводит к большим значениям выделяющейся в тиристоре так называемой мощности коммутационных потерь. Средняя мощность коммутационных потерь может привести к недопустимому перегреву тиристора при большой частоте переключения.

Рис. 5.10. Зависимости напряжения включения тиристора от скорости увеличения подаваемого на его напряжения с учетом только барьерной емкости коллекторного перехода (кривая *1*) и только барьерных емкостей эмиттерных переходов (кривая *2*)

Включение тиристора путем быстрого увеличения напряжения между основными электродами. При быстром нарастании основного напряжения на тиристоре через него будет проходить емкостный ток, обусловленный наличием барьерных емкостей коллекторного и эмиттерных переходов.

Рассмотрим сначала влияние барьерной емкости коллекторного перехода. Емкостный ток через коллекторный переход $i_C = C_к(du/dt)$. Чем больше скорость изменения основного напряжения на тиристоре, тем больше значение емкостного тока через коллекторный переход. Этот ток, проходя через эмиттерные переходы, вызывает увеличение коэффициентов передачи токов эмиттера транзисторных структур, что приводит к включению

тиристора при основном напряжении, меньшем напряжения включения на постоянном токе $U_{вкл0}$ (рис. 5.10).

Барьерные емкости эмиттерных переходов являются причиной появления емкостных токов через эти переходы при быстром изменении основного напряжения на тиристоре. Емкостные токи не связаны с инжекцией носителей заряда, поэтому с увеличением скорости изменения основного напряжения включение тиристора должно происходить при напряжениях, больших $U_{вкл0}$ (рис. 5.10), если учитывать только барьерные емкости эмиттерных переходов.

Практически барьерная емкость коллекторного перехода сказывается сильнее, так как она шунтирует большое активное сопротивление коллекторного перехода, смещенного в обратном направлении при закрытом состоянии тиристора. Барьерные емкости эмиттерных переходов сами оказываются зашунтированными малыми активными сопротивлениями эмиттерных переходов, смещенных при закрытом состоянии тиристора в прямом направлении. Поэтому напряжение включения тиристора с увеличением скорости нарастания основного напряжения уменьшается.

Однако эффект включения тиристоров при большой скорости нарастания основного напряжения часто оказывается не положительным, а отрицательным свойством, так как может приводить к самопроизвольному включения тиристора, например при подключении источника питания. Эффективным способом ослабления этого эффекта является шунтирование эмиттерного перехода (см. § 5.2).

Кроме рассмотренных трех основных способов включения тиристоров можно еще отметить возможность включения тиристора путем освещения кристалла с тиристорной структурой. Но такой тиристор относится к оптоэлектронным полупроводниковым приборам и поэтому будет рассмотрен в гл. 10.

Выключение тиристоров

Выключение тиристора путем уменьшения тока в цепи основных электродов до значения, меньшего удерживающего тока, или путем разрыва цепи основных электродов. Тиристор будет выключен, т. е. переведен из открытого состояния в закрытое, только после рассасывания неравновесных носителей заряда в базовых областях. Если до окончания процесса выключения вновь подать напряжение между основными электродами тиристора, то он окажется во включенном состоянии. Таким образом, для выключения тиристора необходимо некоторое время.

При выключении тиристора путем разрыва цепи основных электродов рассасывание неравновесных носителей заряда происходит только в результате рекомбинации. Такой способ выключения применяется, когда время выключения тиристора не влияет на работу той или иной схемы.

Выключение тиристора путем изменения полярности анодного напряжения. Для ускорения процесса рассасывания неравновесных носителей заряда, накопленных в базовых областях при прохождении прямого тока через открытый тиристор, необходимо понизить потенциальный барьер коллекторного перехода. Однако коллекторный переход при открытом состоянии тиристора уже был смещен в прямом направлении из-за накопленных неравновесных носителей заряда в базовых областях и, следовательно, имел малое сопротивление. Поэтому на долю коллекторного перехода при переключении тиристора на обратное напряжение приходится очень малая часть всего внешнего напряжения. Из-за малого сопротивления тиристора, находящегося еще в открытом

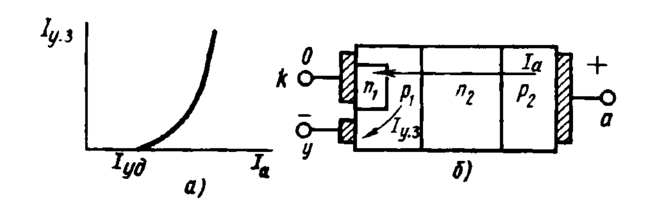


Рис. 5.11. Зависимость тока управления, необходимого для выключения (для запирания) тиристора, от основного тока (*а*) и структура триодного тиристора с направлениями токов при его выключении (*б*)

состоянии, обратный ток на первом этапе процесса выключения ограничен сопротивлением внешней цепи.

Существенное уменьшение времени выключения даже при небольших обратных напряжениях удается получить для тиристоров, проводящих в обратном направлении. У этих тиристоров оба эмиттерные перехода зашунтированы объемными сопротивлениями прилегающих базовых областей (см. § 5.4). Поэтому даже небольшое обратное напряжение способствует быстрому рассасыванию накопленных в базовых областях неравновесных носителей заряда.

Выключение тиристора с помощью тока управляющего электрода. Для выключения тиристора необходимо отвести неравновесные основные носители заряда из базы, у которой имеется управляющий электрод. В то же время основной ток, проходящий через еще открытый тиристор, непрерывно восполняет количество неравновесных носителей заряда в базовых областях. Таким образом, значение тока управления, необходимого для выключения тиристора, зависит от основного тока через тиристор (рис. 5.11, *а*).

Некоторые тиристоры с большой площадью *p-n*-переходов невозможно выключить с помощью тока управляющего электрода при больших токах между основными электродами. Объясняется это тем, что при движении носителей заряда к управляющему

электроду (дырок в тиристоре, структура которого показана на рис. 5.11, *б*) база тиристора под эмиттерным переходом становится неэквипотенциальной и дальние от управляющего электрода части эмиттерного перехода остаются смещенными в прямом направлении. Инжекция электронов из этих частей эмиттерного перехода поддерживает соответствующую часть тиристорной структуры в открытом состоянии. Стремление же еще больше увеличить ток управляющего электрода путем увеличения напряжения на этом электроде может привести к лавинному пробою эмиттерного перехода (в нижней его части по рис. 5.11), так как этот переход имеет обычно малое пробивное напряжение из-за большой концентрации примесей с обеих сторон перехода.

Таким образом, существуют тиристоры, запираемые и незапираемые по управляющему электроду. *Запираемый тиристор* — это тиристор, который может переключаться из закрытого состояния в открытое и наоборот при подаче на управляющий электрод сигналов соответствующей полярности. Но и для запираемого тиристора существует *максимально допустимый постоянный запираемый ток* $I_{з\,max}$ — наибольшее значение основного тока, до которого допускается запирание тиристора по управляющему электроду. При использовании в мощных устройствах запираемые тиристоры обладают преимуществами перед транзисторами, поскольку тиристоры способны выдерживать значительно большие напряжения в закрытом состоянии.

§ 5.7. КОНСТРУКЦИЯ И ТЕХНОЛОГИЯ ИЗГОТОВЛЕНИЯ ТИРИСТОРОВ

Исходный полупроводниковый материал для тиристоров должен иметь большую ширину запрещенной зоны. Тиристоры (как и другие полупроводниковые приборы) на основе широкозонного полупроводника имеют, во-первых, большее значение максимальной рабочей температуры, а следовательно, и максимально допустимой плотности тока в открытом состоянии. Во-вторых, при прочих равных условиях пробивное напряжение выше, что дает возможность делать тиристоры с большим значением напряжения включения и максимально допустимого обратного напряжения. В-третьих, при прочих равных условиях значительно меньший обратный ток через *p-n*-переходы, смещенные в обратном направлении (например, $I_{КБО}$), что обусловливает меньшую мощность, рассеиваемую в тиристоре при закрытом его состоянии и при обратном напряжении. В-четвертых, меньшее значение коэффициента передачи тока эмиттера транзисторных структур, составляющих тиристор, при малых токах (сильнее влияет рекомбинация носителей заряда в эмиттерных переходах) обеспечивает резкую зависимость суммарного коэффициента передачи тока тиристора и более жесткую характеристику его переключения.

Поэтому в массовом производстве тиристоры делают пока только из кремния. Ведутся разработки тиристоров из арсенида галлия и других соединений, имеющих большую ширину запрещенной зоны. Основной метод формирования тиристорных структур на основе кремния — последовательная диффузия различных примесей с использованием планарной технологии.

Для уменьшения вероятности поверхностного лавинного пробоя коллекторного перехода, который может уменьшить напряжение включения тиристора, по периметру кристалла с тиристорной структурой снимают фаску (см. § 3.19). Если при этом высокоомная базовая область имеет толщину, достаточную для устранения эффекта смыкания переходов (см. § 5.1), то напряжение включения таких кремниевых тиристоров может достигать 2500 В.

Площадь *p-n*-переходов рассчитывают, исходя из максимально допустимой плотности тока в статическом режиме через открытый тиристор 200 А/см^2. Соответственно максимально допустимые токи в открытом состоянии для различных тиристоров имеют значения от 40 мА до 1000 А. При этом напряжение в открытом состоянии обычно не превосходит 2 В.

Во время процесса включения триодного тиристора с помощью тока управляющего электрода сначала в открытое состояние переходит только часть тиристорной структуры, расположенной непосредственно около управляющего электрода. Со временем открытое состояние распространяется на всю площадь структуры. Ясно, что при высокой скорости нарастания тока может произойти локальный разогрев структуры до температуры, превышающей максимально допустимую. *Максимально допустимая скорость нарастания тока в открытом состоянии тиристора* $|dI_{\text{откр}}/dt|_{max}$, при которой обеспечивается заданная надежность, может быть повышена в тиристорах с распределенным по всей площади управляющим электродом.

Для уменьшения времени выключения тиристоров, т. е. для ускорения рекомбинации неравновесных носителей заряда, накопленных в базовых областях, необходимо уменьшить время жизни неосновных носителей заряда в высокоомной толстой базе.

Этого можно достичь введением примесей, которые обеспечивают появление вблизи середины запрещенной зоны энергетической диаграммы полупроводника энергетических уровней рекомбинационных ловушек. Для кремния такой примесью, в частности, является золото. Однако с уменьшением времени жизни неосновных носителей заряда в толстой высокоомной базе увеличивается падение напряжения на ней и на тиристоре в открытом состоянии. Уменьшение же толщины базы приводит к уменьшению напряжения включения. Поэтому при конструировании тиристоров приходится искать компромиссное решение.

Контрольные вопросы

1. Что такое тиристор?

2. Почему коллекторный переход тиристора оказывается смещенным в прямом направлении при переключении тиристора из закрытого состояния в открытое?

3. Какие физические явления вызывают увеличение коэффициентов передачи тока эмиттера транзисторных структур, составляющих тиристор?

4. С какой целью в структуре тиристора осуществляют шунтирование одного из эмиттерных переходов, а одну из базовых областей делают толстой по сравнению с диффузионной длиной неосновных носителей заряда?

5. Почему для изготовления тиристора целесообразно использовать полупроводниковый материал с большой шириной запрещенной зоны?

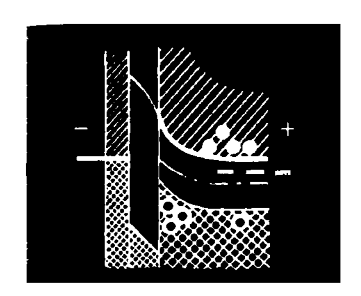
6. В чем преимущества триодного тиристора перед диодным?

7. Какими способами можно перевести тиристор из закрытого состояния в открытое?

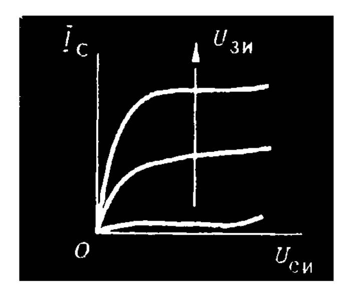
8. Какими способами можно перевести тиристор из открытого состояния в закрытое?

9. Почему не всегда и не все триодные тиристоры можно перевести из открытого состояния в закрытое с помощью тока управления?

10. Какова структура и принцип действия симметричных тиристоров?



глава



Полевые транзисторы и приборы с зарядовой связью

Полевой транзистор — это полупроводниковый прибор, усилительные свойства которого обусловлены потоком основных носителей, протекающим через проводящий канал и управляемым электрическим полем.

§ 6.1. ПОЛЕВЫЕ ТРАНЗИСТОРЫ С УПРАВЛЯЮЩИМ ПЕРЕХОДОМ

Структура и принцип действия

Полевой транзистор с управляющим переходом — это полевой транзистор, управление потоком основных носителей в котором происходит с помощью выпрямляющего электрического перехода, смещенного в обратном направлении.

В качестве выпрямляющего электрического перехода, с помощью которого производят управление потоком основных носителей заряда в полевом транзисторе, может быть *p-n*-переход, гетеропереход или выпрямляющий переход Шотки. Полевые транзисторы с управляющим гетеропереходом делают в основном методом эпитаксии соединений типа $A^{III}B^{V}$ на кристаллы арсенида галлия. Выпрямляющие переходы Шотки осуществляют нанесением металла на кристаллы кремния, арсенида галлия или на различные гетероструктуры. Однако наиболее распространены пока полевые транзисторы с управляющим *p-n*-переходом в кристаллах кремния. Поэтому в дальнейшем в качестве примера рассмотрим полевые транзисторы с управляющим *p-n*-переходом, так как принцип действия полевых транзисторов с различными управляющими переходами идентичен.

Полевой транзистор с управляющим *p-n*-переходом имеет

два омических перехода к области полупроводника, по которой проходит управляемый или регулируемый поток основных носителей заряда, и один или два управляющих *p-n*-перехода, смещенных в обратном направлении (рис. 6.1). При изменении обратного напряжения на управляющем *p-n*-переходе изменяется его толщина, а следовательно, толщина области, по кото-

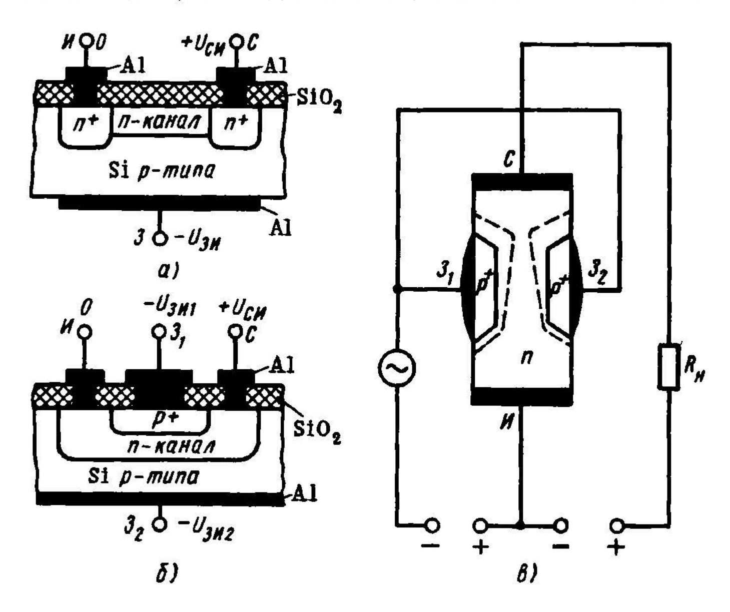


Рис. 6.1. Структуры полевых транзисторов с одним управляющим *p-n*-переходом (*а*), с двумя управляющими *p-n*-переходами (*б*, *в*) и принципиальная схема включения с общим истоком (*в*)

рой проходит управляемый поток основных носителей заряда. Область в полупроводнике, в которой регулируется поток основных носителей заряда, называют *проводящим каналом*. Электрод полевого транзистора, через который в проводящий канал входят носители заряда, называют *истоком*. Электрод полевого транзистора, через который из канала выходят носители заряда, называют *стоком*. Электрод полевого транзистора, на который подают сигнал, называют *затвором*.

Проводящий канал может иметь электропроводность как *n*-, так и *p*-типа. Соответственно различают полевые транзисторы с *n*-каналом и *p*-каналом. Все полярности напряжений смещения, подаваемых на электроды полевых транзисторов с *n*- и *p*-каналом, противоположны. Полевые транзисторы, имеющие проводящие каналы с разными типами электропроводности, называют *комплементарными*, т. е. дополняющими друг друга по типу электропроводности проводящих каналов. Комплементарные транзисторы оказываются удобными элементами для построения ряда схем.

Управление током стока, т. е. током от внешнего относительно мощного источника питания в цепи нагрузки, происходит при изменении обратного напряжения на *p-n*-переходе затвора (или на двух *p-n*-переходах одновременно). В связи с малостью обратных токов мощность, необходимая для управления током стока и потребляемая от источника сигнала в цепи затвора, оказывается ничтожно малой. Поэтому полевой транзистор может обеспечить усиление электрических сигналов как по мощности, так и по току и напряжению.

Таким образом, полевой транзистор по принципу действия аналогичен вакуумному триоду. Исток в полевом тразисторе подобен катоду вакуумного триода, затвор — сетке, сток — аноду. Но при этом полевой транзистор существенно отличается от вакуумного триода. Во-первых, для работы полевого транзистора не требуется подогрева катода. Во-вторых, любую из функций истока и стока может выполнять каждый из этих электродов. В-третьих, полевые транзисторы могут быть сделаны как с *n*-каналом, так и с *p*-каналом, что позволяет удачно сочетать эти два типа полевых транзисторов в схемах.

От биполярного транзистора и в том числе от однопереходного транзистора полевой транзистор отличается, во-первых, принципом действия: в биполярном транзисторе управление выходным сигналом производится входным током, а в полевом транзисторе — входным напряжением или электрическим полем. Во-вторых, полевые транзисторы имеют значительно большие входные сопротивления, что связано с обратным смещением *p-n*-перехода затвора в рассматриваемом типе полевых транзисторов. В-третьих, полевые транзисторы могут обладать низким уровнем шума (особенно на низких частотах), так как в полевых транзисторах не используется явление инжекции неосновных носителей заряда и канал полевого транзистора отделен от поверхности полупроводникового кристалла (рис. 6.1, *б, в*). Процессы рекомбинации носителей в *p-n*-переходе и в базе биполярного транзистора, а также генерационно-рекомбинационные процессы на поверхности кристалла полупроводника сопровождаются возникновением низкочастотных шумов.

Статические характеристики

Полевые транзисторы имеют большие входные и выходные сопротивления. Поэтому в отличие от биполярных транзисторов статические характеристики полевых транзисторов удобнее и правильнее исследовать с помощью источников напряжения в качестве источников питания. Следовательно, статические характеристики полевых транзисторов должны соответствовать системе уравнений

$$\left.\begin{aligned} I_1 &= f_1(U_1;\ U_2)\ ; \\ I_2 &= f_2(U_1;\ U_2)\ . \end{aligned}\right\} \qquad (6.1)$$

Важнейшими семействами статических характеристик для полевого транзистора являются семейство выходных статических характеристик и семейство статических характеристик передачи.

Выходные статические характеристики полевого транзистора представляют собой зависимости тока стока от напряжения на стоке относительно истока при различных постоянных напряжениях на затворе (рис. 6.2, *а*). Рассмотрим вначале характер одной зависимости $I_С = f(U_{СИ})$ при $U_{ЗИ} = 0$. Напряжение на затворе относительно истока будет равно нулю только в том случае,

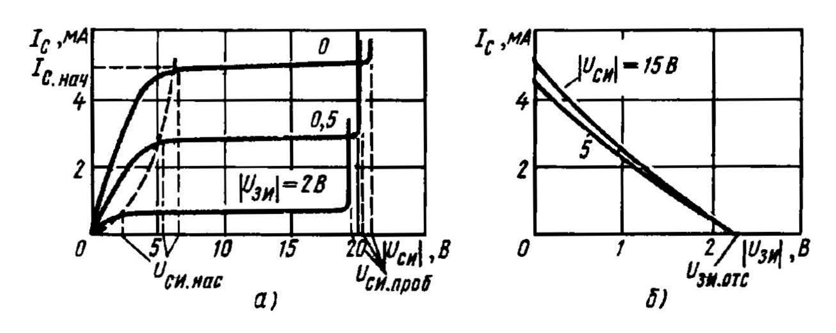


Рис. 6.2. Выходные статические характеристики (*а*) и статические характеристики передачи (*б*) полевого транзистора с управляющим *p-n*-переходом

если затвор закорочен с истоком. Характеристика выходит из начала координат под углом, соответствующим начальному статическому сопротивлению канала и сопротивлениям $R_И$ и $R_С$ прилегающих к каналу областей полупроводникового кристалла с тем же типом электропроводности. Статнческое сопротивленне канала определяется его длиной и поперечным сечением, зависящим от толщины *p-n*-перехода (или *p-n*-переходов).

Первая часть характернстики, которую называют *крутой частью,* сублинейна, т. е. ток стока растет замедленно с ростом напряжения на стоке. Объясняется эта нелинейность характеристики увеличением толщины *p-n*-перехода затвора около стока, так как с увелнчением напряжения на стоке растет по абсолютному значению обратное напряжение на *p-n*-переходе затвора. Ток стока, проходя по каналу, создает его неэквипотенциальность. Таким образом, наибольшая толщина *p-n*-перехода и соответственно наименьшее поперечное сеченне канала получаются со стороны стока (см. рнс. 6.1, *в* и 6.3).

Другой физической причиной, приводящей к сублинейности выходной характеристики, является уменьшение подвижности носителей заряда в канале при увеличении в нем напряженности электрического поля (см. § 1.10).

Прн некотором напряжении на стоке $U_{СИ нас}$ — *напряжении насыщения* — происходит перекрытие канала из-за увеличения

толщины p-n-перехода затвора. Ток стока при дальнейшем увеличеиии напряжения на стоке почти не растет.

При напряжении между затвором и истоком, равном иулю, и прн напряжении на стоке, равном или превышающем напряжение насыщения, ток стока называют *начальным током стока* $I_{С нач}$. Часть характеристики, соответствующую насыщению тока стока, называют *пологой частью*. Следует учитывать условность понятия «перекрытие» каиала при увеличении напряжения на стоке и неизменном напряжении на затворе относительно истока, так как перекрытие канала прн указанных условиях является следствием увеличения тока стока. Таким образом, можно считать, что в результате увеличения тока стока или напряжения на стоке автоматически устанавливается некоторое малое сечеиие канала со стороны стокового электрода.

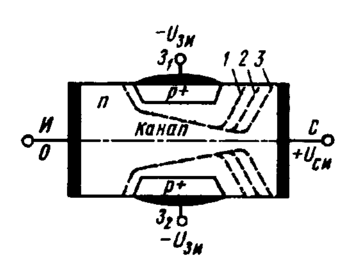


Рис. 6.3. Увеличение толщины управляющих p-n-переходов со стороны стока и соответствующее увеличение длины перекрытой части проводящего канала при увеличении напряжения на стоке и неизменном напряжении на затворе относительно истока:

1 — напряжение на стоке равно напряжению насыщения $U_{СИ\,нас}$; 2 — $U'_{СИ} > U_{СИ\,нас}$; 3 — $U''_{СИ} > U'_{СИ}$

При дальнейшем увеличении напряжения на стоке увеличивается длина перекрытой части канала и растет статическое сопротивление канала. Если бы длина перекрытой части канала увеличивалась пропорционально напряжению на стоке, то ток стока не изменялся бы при напряжениях на стоке, превышающих напряжение насыщения. Однако длина перекрытой части канала увеличивается из-за увеличения толщины p-n-перехода с ростом напряжения на стоке (рис. 6.3), а толщина p-n-перехода пропорциональна либо корню квадратному, либо корню кубическому из напряжения (см. § 2.6 и 2.7). Поэтому в пологой части характеристики наблюдается некоторое увеличение тока стока при увеличении напряжения на стоке.

Теперь рассмотрим смещение и изменение статических характеристик с изменением напряжеиия на затворе. При подаче на затвор напряжения такой полярности относительно истока, которая соответствует обратному смещению p-n-перехода затвора, и при увеличении этого напряжения по абсолютному значению уменьшается начальиое поперечиое сечение канала. Поэтому начальные участки выходиых статических характеристик при напряжениях на затворе, отличных от нуля, имеют другой наклон, соответствующий бóльшим начальным статическим сопротивлениям канала.

При меньших начальных поперечиых сечениях перекрытие канала из-за увеличения напряжения на стоке происходит при меньших напряжениях насыщения (см. рис. 6.2, *а*).

При больших напряжениях на стоке может возникнуть пробой

p-n-перехода затвора. Обратное напряжеиие на *p-n*-переходе затвора измеияется вдоль длины канала, достигая максимальиого значеиия у стокового конца канала. Напряжение, приложенное к *p-n*-переходу затвора в этом месте, является суммой напряжеиий на стоке и на затворе. Таким образом, пробой полевого транзистора может происходить при разных напряжеииях на стоке в зависнмости от иапряжеиия иа затворе. Чем больше иапряжение иа затворе, тем меиьше напряжение на стоке, при котором произойдет пробой *p-n*-перехода затвора (см. рис. 6.2, *а*). Полевые транзисторы делают обычио иа основе кремиия. Поэтому пробой таких траизисторов имеет лавинный характер.

Статические характеристики передачи полевого транзистора в соответствии с (6.1) представляют собой зависимости тока стока от напряжения на затворе при различиых постояиных иапряжениях на стоке. Так как основным рабочим режимом полевых транзисторов является режим насыщеиия тока стока, что соответствует пологим частям выходиых статических характеристик, то иаибольший интерес представляет зависимость тока насыщения от иапряжения на затворе при постоянном иапряжеиии на стоке. Характер этой зависимости ясен из принципа действия полевого траизистора с управляющим *p-n*-переходом. При изменении напряжеиия иа стоке смещением характеристик передачи практически можио пренебречь в связи с малым изменением тока в пологой части выходиых статических характеристик (см. рис. 6.2, *б*).

Напряжеиие между затвором и истоком полевого траизистора с управляющим переходом, при котором ток стока достигает задаииого иизкого зиачеиия, иазывают *напряжением отсечки полевого транзистора* $U_{ЗИ\,отс}$.

При рассмотреиии статических характеристик полевого траизистора были отмечеиы его осиовиые статические параметры. По статической характеристике передачи можио определить еще одии осиовной параметр полевого транзистора, характеризующий его усилительиые свойства, — *крутизну характеристики* полевого траизистора S, которая представляет собой отиошение измеиеиия тока стока к изменению напряжеиия иа затворе при коротком замыкании по перемениому току на выходе траизистора в схеме с общим истоком:

$$S = \left.\frac{dI_С}{dU_{ЗИ}}\right|_{U_{СИ} = \text{const}}.$$

Крутизиа характеристики полевого транзистора составляет обычио иесколько миллиампер на вольт.

§ 6.2. РАСЧЕТ ВЫХОДНЫХ СТАТИЧЕСКИХ ХАРАКТЕРИСТИК ПОЛЕВОГО ТРАНЗИСТОРА С УПРАВЛЯЮЩИМ ПЕРЕХОДОМ

Пренебрегая объемными сопротивлениями кристалла полупроводиика на участках между коицами каиала и контактами истока и стока, рабочую часть полевого траизистора можно представить в упрощеином виде (рис. 6.4).

Плотность тока в канале

$$J = \gamma E = -\gamma \frac{d\varphi}{dx}, \qquad (6.2)$$

где γ — удельная проводимость каиала.

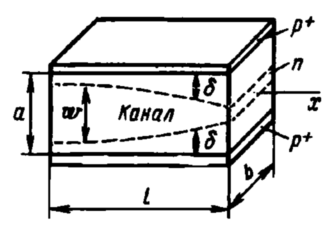


Рис. 6.4. Структура рабочей части полевого транзистора с управляющими p-n-переходами

В первом приближении будем считать удельиую проводимость материала каиала не зависящей от иапряженности электрического поля, т. е. не будем учитывать изменение подвижиости. Плотность тока в канале измеияется по его длиие в связи с изменением сечения каиала и соответствующим измеиением напряжеиности.

Ток в каиале траизистора, неизмениый по всему каналу,

$$I_C = -\gamma b w \frac{d\varphi}{dx}, \qquad (6.3)$$

где b — шириниа каиала.

Толщиниа каиала w зависит от толщины p-n-переходов:

$$w = a - 2\delta. \qquad (6.4)$$

Толщина p-n-перехода зависит от напряжеиия иа ием (см. § 2.6). В выражеиии (2.27) для толщиины резкого p-n-перехода можио преиебречь коитактной разиостью потенциала на p-n-переходе по сравиеиию с отиосительио большим обратным напряжеиием, подаиным иа затвор полевого транзистора. Однако необходимо учесть иеэквипотеициальность каиала, возиикающую из-за прохождеиия тока по каиалу от истока к стоку. Тогда

$$\delta = \sqrt{\frac{2\varepsilon\varepsilon_0}{qN}(\varphi - U_{ЗИ})}. \qquad (6.5)$$

Для более компактиой записи иайдем напряжеиие отсечки $U_{ЗИ\,отс}$. Из определеиия иапряжеиия отсечки следует, что толщина каиала около стока при этом напряжеиии равиа иулю, а толщина p-n-перехода

$$\delta = \frac{a}{2} = \sqrt{\frac{2\varepsilon\varepsilon_0}{qN}} U_{ЗИ\,отс}^{1/2}. \qquad (6.6)$$

Используя (6.4), (6.5) и (6.6), получим

$$w = a\left(1 - \sqrt{\frac{\varphi + U_{\text{ЗИ}}}{U_{\text{ЗИ отс}}}}\right). \qquad (6.7)$$

После подстановки (6.7) в (6.3) абсолютное значение тока стока

$$|I_{\text{С}}| = \gamma b a\left(1 - \sqrt{\frac{\varphi + U_{\text{ЗИ}}}{U_{\text{ЗИ отс}}}}\right)\frac{d\varphi}{dx}. \qquad (6.8)$$

Решая это дифференциальное уравнение с разделяющимися переменными при граничиых условиях:

1) $x = 0$, $\varphi = 0$ и 2) $x = l$, $\varphi = U_{\text{СИ}}$,

получим искомую функцию

$$I_{\text{С}} = \frac{1}{R_{\text{СИ отк}}}\left[U_{\text{СИ}} + \frac{2}{3}\,\frac{U_{\text{ЗИ}}^{3/2} - (U_{\text{СИ}} + U_{\text{ЗИ}})^{3/2}}{U_{\text{ЗИ отс}}^{1/2}}\right], \qquad (6.9)$$

где $R_{\text{СИ отк}} = l/(\gamma a b)$ — сопротивление сток — исток в открытом состоянии, т. е. при $U_{\text{ЗИ}} = 0$ и при малом напряжении на стоке (меньшем напряжения насыщения).

Выражение (6.9) дает возможность иайти ток насыщення полевого траизистора. Как отмечалось, перекрытие канала происходит при иапряжеинии отсечки $U_{\text{ЗИ отс}}$. Режим иасыщения наступит при условии

$$U_{\text{ЗИ}} + U_{\text{СИ}} = U_{\text{ЗИ отс}},$$

т. е. при напряжении иа стоке

$$U_{\text{СИ нас}} = U_{\text{ЗИ отс}} - U_{\text{ЗИ}}. \qquad (6.10)$$

Если в выражеиии (6.9) замеиить напряжение на затворе $U_{\text{ЗИ}}$ напряжением на стоке, при котором происходит насыщение, и иапряжеиием отсечки $U_{\text{ЗИ отс}}$, пользуясь соотиошением (6.10), то получим связь между током и напряжением иасыщеиия:

$$I_{\text{С нас}} = \frac{1}{R_{\text{СИ отк}}}\left[U_{\text{СИ нас}} + \frac{2}{3}(U_{\text{ЗИ отс}} - U_{\text{СИ нас}}) \times \right.$$

$$\left. \times \sqrt{\frac{U_{\text{ЗИ отс}} - U_{\text{СИ нас}}}{U_{\text{ЗИ отс}}}} - \frac{2}{3}U_{\text{ЗИ отс}}\right]. \qquad (6.11)$$

На рис. 6.2 штриховая кривая показывает зависимость $I_{\text{С нас}} = f(U_{\text{СИ нас}})$. Эта зависимость является геометрическим местом точек, соответствующих токам и напряжениям, при которых наступает режим насыщения полевого траизистора.

Важио знать также зависимость тока насыщения от иапряжеиия иа затворе полевого транзистора, т. е. характеристику передачи. Эту зависимость можно найти, если в выражение (6.9) подставить из условия насыщения (6.10) напряжение иа стоке $U_{\text{СИ нас}}$:

$$I_{\text{Снас}} = \frac{1}{R_{\text{СИотк}}} \left[\frac{1}{3} U_{\text{ЗИотс}} - U_{\text{ЗИ}} \left(1 - \frac{2}{3} \sqrt{\frac{U_{\text{ЗИ}}}{U_{\text{ЗИотс}}}} \right) \right] \qquad (6.12)$$

Поскольку пологий участок выходиых статических характеристик полевых транзисторов является, как и в вакуумных пеитодах, основным рабочим участком, определим крутизну характеристики S именно в этой области. Диффереицируя (6.12) по $U_{\text{ЗИ}}$, получим

$$S = \left| \frac{dI_{\text{Снас}}}{dU_{\text{ЗИ}}} \right| = \frac{1}{R_{\text{СИотк}}} \left(1 - \sqrt{\frac{U_{\text{ЗИ}}}{U_{\text{ЗИотс}}}} \right). \qquad (6.13)$$

Из выражения (6.13) следует, что для получеиия больших значеиий этого параметра иеобходимо иметь меиьшее значение сопротивления сток — исток в открытом состоянии транзистора $R_{\text{СИотк}}$ или бóльшую удельную проводимость исходного материала. В то же время концентрация примесей и соответственно носителей заряда в канале должиа быть иебольшой, чтобы при увеличеиии напряжения на *p-n*-переходе ои расширялся в сторону каиала. Таким образом, для получения больших зиачений крутизны характеристики желательно при изготовлеиии полевого транзистора выбирать материал с большей подвижиостью носителей заряда.

Значение крутизны характеристики полевого транзистора в значительной мере определяется отношеиием ширины канала b к его длиие l. Увеличение отношения b/l позволяет повысить крутизну характеристики S и ток насыщения полевого транзистора [см. (6.11) и (6.13)].

Согласно выраженню (6.13), для повышения крутизны характеристики S необходимо увеличивать толщииу каиала a. Однако с увеличением толщииы канала недопустимо растет иапряжеиие отсечки и напряжеиие насыщения, соответствующее входу полевого транзистора в режим насыщения. Так как режим иасыщеиия является осиовным рабочим режимом полевого транзистора, напряжение отсечки должно быть малым. Поэтому толщииу канала стараются сделать небольшой, несмотря на некоторое уменьшение крутизны характеристики.

§ 6.3. ЭКВИВАЛЕНТНЫЕ СХЕМЫ ПОЛЕВОГО ТРАНЗИСТОРА С УПРАВЛЯЮЩИМ ПЕРЕХОДОМ

Исходя из прииципа действия и структуры полевого транзистора, можно составить его эквивалентную схему для низких частот (рис. 6.5).

Сопротивления $r_{\text{с}}$ и $r_{\text{и}}$ представляют собой объемиые сопротивления кристалла полупроводника на участках между концами канала и контактами стока и истока соответствеино. Эти сопротивления зависят от конструкции транзистора и технологии его изготовления. На низких частотах влиянием сопротивления $r_{\text{с}}$

можно пренебречь по сравнению с обычно большим сопротивлением нагрузки в цепи стока и большим дифференциальным сопротивлением канала r_i. Общее для входной и выходной цепей сопротивление $r_и$ является сопротивлением внутренней обратной связи в полевом транзисторе, включенном по схеме с общим истоком. Падение напряжения на этом сопротивлении при прохождении тока стока оказывается обратным для p-n-перехода. В свою очередь, увеличение обратного напряжения на p-n-переходе затвора транзистора приводит к уменьшению тока стока.

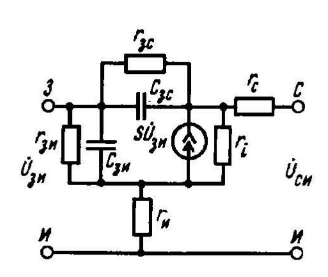


Рис. 6.5. Физическая эквивалентная схема полевого транзистора

Емкости $C_{зи}$ и $C_{зс}$, сопротивления $r_{зи}$ и $r_{зс}$ замещают в этой эквивалентной схеме p-n-переход с его барьерной емкостью и большим активным дифференциальным сопротивлением при обратном смещении.

Генератор тока, включенный параллельно сопротивлению канала, отражает усилительные свойства транзистора. Ток этого генератора пропорционален входному напряжению $\dot{U}_{зи}$; коэффициентом пропорциональности является крутизна характеристики S.

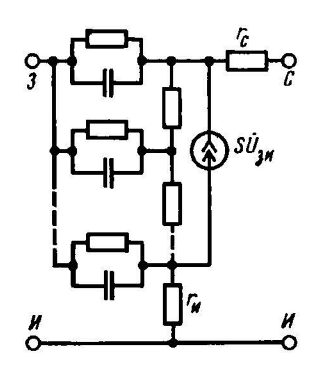


Рис. 6.6. Физическая эквивалентная схема с распределенными параметрами полевого транзистора

Нужно учитывать, что емкость и сопротивление затвора распределены по всей его площади и что сопротивление канала также является распределенным. В этом случае эквивалентная схема полевого транзистора должна быть представлена в виде схемы с распределенными параметрами (рис. 6.6). Однако такая схема значительно сложнее для выяснения свойств и характеристик полевых транзисторов.

Кроме физических эквивалентных схем полевого транзистора можно представить и формальные эквивалентные схемы с y-, z- или h-параметрами. Так как входные и выходные сопротивления полевых транзисторов велики, то удобнее измерять и задавать комплексные параметры проводимостей его формальной эквивалентной схемы (рис. 6.7). Токи и напряжения на выводах полевого транзистора в режиме малого сигнала для схемы с общим истоком соответствуют следующим характеристическим уравнениям четырехполюсника:

$$\left.\begin{aligned} \dot{I}_з &= y_{11и}\dot{U}_{зи} + y_{12и}\dot{U}_{си}\,; \\ \dot{I}_с &= y_{21и}\dot{U}_{зи} + y_{22и}\dot{U}_{си}\,. \end{aligned}\right\} \qquad (6.14)$$

Определяются y-параметры при режимах короткого замыкания по переменному току на выходе и входе транзистора:

$$\left.\begin{aligned} y_{11и} &= \frac{\dot{I}_з}{\dot{U}_{зи}}\bigg|_{\dot{U}_{си}=0}; \quad y_{12и} = \frac{\dot{I}_з}{\dot{U}_{си}}\bigg|_{\dot{U}_{зи}=0}; \\ y_{21и} &= \frac{\dot{I}_с}{\dot{U}_{зи}}\bigg|_{\dot{U}_{си}=0}; \quad y_{22и} = \frac{\dot{I}_с}{\dot{U}_{си}}\bigg|_{\dot{U}_{зи}=0}. \end{aligned}\right\} \quad (6.15)$$

Если эти режимы воспроизвести на эквивалентной схеме рис. 6.5, то можно найти формулы перехода от параметров сосредоточенных элементов физической эквивалеитиой схемы к y-пара-

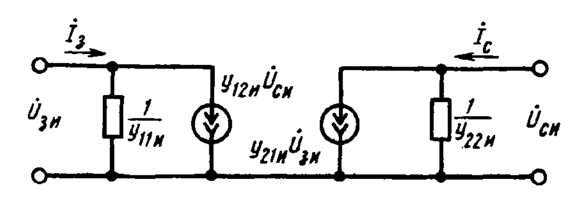


Рис. 6.7. Формальная эквивалентная схема полевого транзистора, соответствующая системе y-параметров

метрам. Пренебрегая малыми проводимостями p-n-перехода затвора и объемными сопротивлениями полупроводника около истока и стока, получим

$$\left.\begin{aligned} y_{11и} &= j\omega(C_{зс} + C_{зи}); \quad y_{12и} = -j\omega C_{зс}; \\ y_{21и} &= S - j\omega C_{зи}; \quad y_{22и} = \frac{1}{r_i} + j\omega C_{зс}. \end{aligned}\right\} \quad (6.16)$$

Все эти параметры зависят от значений постоянных смещений на электродах полевого транзистора.

§ 6.4. ЧАСТОТНЫЕ СВОЙСТВА ПОЛЕВЫХ ТРАНЗИСТОРОВ С УПРАВЛЯЮЩИМ ПЕРЕХОДОМ

Принцип действия полевого транзистора не связан с инжекцией неосновных носителей заряда в базу и их относительно медленным движением до коллекторного перехода. Полевой транзистор — это прибор без инжекции. Поэтому иенрционность и частотные свойства полевого траизистора с управляющим переходом обусловлеиы инерционностью процесса заряда и разряда барьерной емкости p-n-перехода затвора. Напряжение на затворе измеииться мгиовенно не может, так как барьерная емкость p-n-перехода затвора перезаряжается токами, проходящими через распределенное сопротивление канала и через объемные сопротивления кри-

сталла полупроводника около истока и стока. Поэтому не может мгновенно измениться и сечение канала.

На низких частотах полное входное сопротивление полевого транзистора с управляющим *p-n*-переходом определяется большим значением $r_{зи}$. С ростом частоты входного сигнала полное входное сопротивление транзистора уменьшается в связи с наличием емкости $C_{зи}$. Следовательно, для управления полевым транзистором при высоких частотах необходима большая мощность входного сигнала.

Кроме того, наличие в полевом транзисторе проходной емкости $C_{зс}$, аналогичной емкости анод — сетка в вакуумной лампе, приводит к возникновению в полевом транзисторе частотно-зависимой обратной связи. С ростом частоты увеличивается обратная связь через цепь $r_{с}C_{зс}$ (см. рис. 6.5), что эквивалентно уменьшению полного входного сопротивления полевого транзистора и уменьшению его усиления.

Для получения оптимального усиления в реальных схемах на полевых транзисторах необходимо согласование внешних сопротивлений с входным и выходным сопротивлениями транзистора. Поэтому во внешней цепи входа и выхода полевого транзистора обычно есть большие сопротивления, которые значительно увеличивают постоянные времени перезаряда емкостей полевого транзистора.

В связи с перечисленными причинами максимальные рабочие частоты реальных схем на полевых транзисторах с управляющим *p-n*-переходом не превышают нескольких сотен мегагерц.

§ 6.5. ПОЛЕВЫЕ ТРАНЗИСТОРЫ С ИЗОЛИРОВАННЫМ ЗАТВОРОМ

Полевой транзистор с изолированным затвором — это полевой транзистор, имеющий один или несколько затворов, электрически изолированных от проводящего канала.

Структуры полевых транзисторов с изолированным затвором показаны на рис. 6.8. В кристалле полупроводника с относительно высоким удельным сопротивлением, который называют *подложкой,* созданы две сильнолегированные области с противоположным типом электропроводности. На эти области нанесены металлические электроды — *исток* и *сток.* Расстояние между сильнолегированными областями истока и стока может составлять всего несколько микрометров. Поверхность кристалла полупроводника между истоком и стоком покрыта тонким (порядка 0,1 мкм) слоем диэлектрика. На слой диэлектрика нанесен металлический электрод — *затвор.* Получается структура, состоящая из слоя металла, диэлектрика и полупроводника, т. е. *МДП-структура.*

Полевой транзистор с изолированным затвором, в котором в качестве изоляционного слоя между металлическим затвором и проводящим каналом использован диэлектрик, называют полевым

транзистором типа металл — диэлектрик — полупроводник или *МДП-транзистором*.

Выпрямляющие электрические переходы под истоком и стоком могут быть выполнены не только в виде *p-n*-переходов, но и в виде выпрямляющих переходов Шотки, т. е. путем нанесения металлических электродов истока и стока непосредственно на подложку. Использование выпрямляющих переходов Шотки под истоком и стоком может обеспечить ряд преимуществ в технологии изготов-

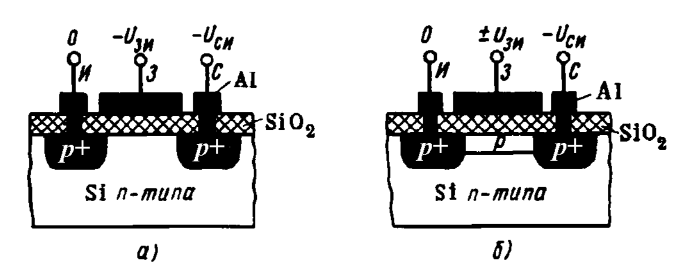


Рис. 6.8. Структуры полевых транзисторов с изолированным затвором (с *p*-каналом):
а — с индуцированным каналом; *б* — со встроенным каналом

ления таких транзисторов, а также улучшит их характеристики. Так, переходы Шотки имеют сравнительно малую толщину, что облегчает создание полевых транзисторов с очень коротким проводящим каналом. Формирование переходов Шотки происходит при относительно низкой температуре, а исключение высокотемпературных операций способствует повышению качества диэлектрического слоя, а также минимизации размеров транзисторной структуры. Кроме того, МДП-транзисторы с переходами Шотки под истоком и стоком могут быть изготовлены на полупроводниках, в которых не удается получить достаточно качественных *p-n*-переходов.

Пока исходным полупроводником для полевых транзисторов с изолированным затвором в основном является кремний. Поэтому в качестве диэлектрика под затвором используется обычно слой диоксида кремния SiO_2, выращенный на поверхности кристалла кремния путем высокотемпературного окисления.

Полевой транзистор с изолированным затвором, в котором в качестве изоляционного слоя между каждым металлическим затвором и проводящим каналом использован оксид полупроводника, называют полевым транзистором типа металл — оксид — полупроводник или *МОП-транзистором*.

Однако для этих транзисторов чаще используют более общий термин — МДП-транзисторы.

Существуют две разновидности МДП-транзисторов: с индуцированным каналом и со встроенным каналом.

В МДП-транзисторах *с индуцированным каналом* (рис. 6.8, *а*)

проводящий канал между сильнолегированными областями истока и стока и, следовательно, заметный ток стока появляются только при определенной полярности и при определенном значении напряжения на затворе относительно истока, которое называют *пороговым напряжением* ($U_{\text{ЗИ пор}}$).

В МДП-транзисторах *со встроенным каналом* (рис. 6.8, *б*) у поверхности полупроводника под затвором при нулевом напряжении на затворе относительно истока существует инверсный слой — канал, который соединяет исток со стоком.

Изображенные на рис. 6.8 структуры полевых транзисторов с изолированным затвором имеют подложку с электропроводностью *n*-типа. Поэтому сильнолегированные области под истоком и стоком, а также индуцированный и встроенный канал имеют электропроводность *p*-типа. Если же аналогичные транзисторы созданы на подложке с электропроводностью *p*-типа, то канал у них будет иметь электропроводность *n*-типа.

МДП-транзисторы с индуцированным каналом

Принцип действия. При напряжении на затворе относительно истока, равном нулю, и при наличии напряжения на стоке ток стока оказывается ничтожно малым. Он представляет собой обратный ток *p-n*-перехода между подложкой и сильнолегированной областью стока. При отрицательном потенциале на затворе (для структуры, показанной на рис. 6.8, *а*) в результате проникновения электрического поля через диэлектрический слой в полупроводник при малых напряжениях на затворе (меньших $U_{\text{ЗИ пор}}$) у поверхности полупроводника под затвором возникают объединенный основными носителями заряда слой и область объемного заряда, состоящая из ионизированных нескомпенсированных примесных атомов. При напряжениях на затворе, бóльших порогового $U_{\text{ЗИ пор}}$, у поверхности полупроводника под затвором возникает инверсный слой, который и является проводящим каналом между истоком и стоком. С изменением напряжения на затворе изменяется концентрация носителей заряда в проводящем канале, а также толщина или поперечное сечение проводящего канала, т. е. происходит модуляция сопротивления проводящего канала. Основной причиной модуляции сопротивления проводящего канала в МДП-транзисторах с индуцированным каналом является изменение концентрации носителей заряда в проводящем канале; в полевых транзисторах с управляющим переходом — изменение толщины или поперечного сечения канала.

При изменении сопротивления проводящего канала изменяется и ток стока (схема включения МДП-транзистора аналогична схеме включения полевого транзистора с управляющим переходом, показанной на рис. 6.1, *в*, где полярности источников питания зависят от типа электропроводности проводящего канала). Так

происходит управление током стока в МДП-транзисторе с индуцированным каналом.

В связи с тем что затвор отделен от подложки диэлектрическим слоем, ток в цепи затвора ничтожно мал, мала и мощность, потребляемая от источника сигнала в цепи затвора и необходимая для управления относительно большим током стока. Таким образом, МДП-транзистор с индуцированным каналом может производить усиление электрических сигналов по напряжению и по мощности.

Принцип усиления мощности в МДП-транзисторах можно рассматривать с точки зрения передачи носителями заряда энергии постоянного электрического поля (энергии источника питания в выходной цепи) переменному электрическому полю. Этот общий для различных приборов принцип усиления мощности был рассмотрен в § 4.2 для объяснения усиления по мощности биполяр-

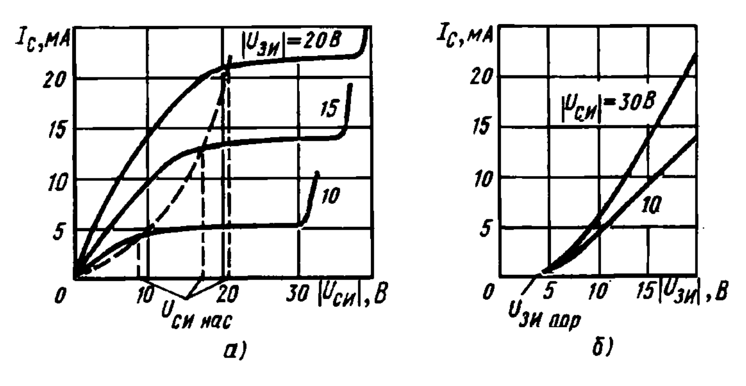


Рис. 6.9. Выходные статические характеристики (*а*) и статические характеристики передачи (*б*) МДП-транзистора с индуцированным каналом

ных транзисторов. В МДП-транзисторе до возникновения канала почти все напряжение источника питания в цепи стока падало на полупроводнике между истоком и стоком, создавая относительно большую постоянную составляющую напряженности электрического поля. Под действием напряжения на затворе в полупроводнике под затвором возникает канал, по которому от истока к стоку движутся носители заряда — дырки. Дырки, двигаясь по направлению постоянной составляющей электрического поля, разгоняются этим полем, и их энергия увеличивается за счет энергии источника питания в цепи стока. Одновременно с возникновением канала и появлением в нем подвижных носителей заряда уменьшается напряжение на стоке, т. е. мгновенное значение переменной составляющей электрического поля в канале направлено противоположно постоянной составляющей. Поэтому дырки тормозятся переменным электрическим полем, отдавая ему часть своей энергии.

Выходные статические характеристики. Характер зависимостей $I_С = f(U_{СИ})$ при $U_{ЗИ} = \text{const}$ для МДП-транзистора с индуцированным каналом аналогичен характеру таких же зависимостей для полевого транзистора с управляющим переходом (см. § 6.1). Сублинейность крутых частей характеристик (рис. 6.9, *а*) объясняется уменьшением толщины канала около стока при увеличении напряжения на стоке и неизменном напряжении на затворе, так как на сток и на затвор подаются потенциалы одного знака относительно истока. Следовательно, разность потенциалов между стоком и затвором или между затвором и прилегающей к стоку частью канала уменьшается. Другими словами, из-за прохождения по каналу тока стока получается неэквипотенциальность канала по его длине. Поэтому при увеличении тока стока происходит уменьшение поперечного сечения канала около стока.

При напряжении насыщения $U_{СИ\,нас}$ происходит перекрытие канала около стока и дальнейшее увеличение напряжения на стоке вызывает очень малое увеличение тока стока. Распределение напряженности электрического поля у поверхности полупроводника при напряжении на стоке, превышающем напряжение насыщения, т. е. для пологой части выходных статических характеристик, показано на рис. 6.10. На расстоянии l_1 от сильнолегированной области истока преобладает нормальная составляющая напряженности электрического поля, созданная напряжением на затворе. На этом участке существует инверсный слой у поверхности полупроводника. На расстоянии l_2 от сильнолегированной области стока преобладает касательная составляющая электрического поля, созданная напряжением на стоке относительно истока. Несмотря на то, что на участке канала протяженностью l_2 нормальная составляющая напряженности имеет другое направление и отталкивает дырки от поверхности полупроводника, через этот перекрытый участок канала идет ток, связанный с движением дырок под действием сильного тянущего поля (касательной составляющей).

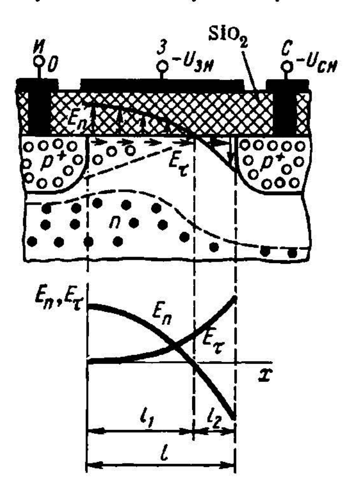


Рис. 6.10. Распределение напряженности электрического поля у поверхности полупроводника под затвором при напряжении на стоке, превышающем напряжение насыщения

Сублинейный характер зависимостей $I_С = f(U_{СИ})$ вызван также эффектом насыщения дрейфовой скорости носителей заряда или уменьшением их подвижности в сильных полях, как и в полевых транзисторах с управляющим *p-n*-переходом.

При увеличении напряжения на затворе (по абсолютному

значению) выходные статические характеристики смещаются в область больших токов стока (см. рис. 6.9, *а*), что легко понять на основе принципа действия МДП-транзистора с индуцированным каналом.

При больших напряжениях на стоке может произойти пробой МДП-транзистора, при этом может быть два вида пробоя: пробой *р-п*-перехода под стоком и пробой диэлектрика под затвором.

Пробой *р-п*-перехода обычно имеет лавинный характер, так как МДП-транзисторы изготовляют обычно иа кремнии. При этом на пробивное напряжение $U_{\text{СИ проб}}$ может влиять напряжение на затворе: так как на сток и на затвор МДП-транзистора с индуцированным каналом подаются потенциалы одной полярности, то с увеличением напряжения на затворе будет увеличиваться $U_{\text{СИ проб}}$ (см. рис. 6.9, *а*).

Пробой диэлектрика под затвором может происходить при напряжении на затворе всего в несколько десятков вольт, так как толщина слоя диоксида кремния около 0,1 мкм. Пробой обычно имеет тепловой характер, происходит при шнуровании тока (см. § 3.13), и поэтому даже при небольших энергиях импульсов напряжения могут произойти необратимые изменения в диэлектрике. Этот вид пробоя может возникать в результате накопления статических зарядов, так как входное сопротивление МДП-транзисторов велико. Для исключения возможности такого вида пробоя вход МДП-транзистора часто защищают стабилитроном, ограиичивающим иапряжение на затворе.

Статические характеристики передачи. Характер зависимостей $I_{\text{С}}=f(U_{\text{ЗИ}})$ при $U_{\text{СИ}}=\text{const}$ ясен из принципа действия МДП-транзистора с индуцированным каналом. Характеристики для разных напряжений на стоке выходят из точки на оси абсцисс, соответствующей пороговому напряжению $U_{\text{ЗИ пор}}$ (см. рис. 6.9,*б*). С увеличением напряжения на стоке при неизменном напряжении на затворе ток стока возрастает даже в пологой части статических выходных характеристик (см. рис. 6.9, *а*), что приводит к смещению характеристик передачи вверх в выбранной системе координат.

Интересным и важным с точки зрения примененияя МДП-транзисторов является температурное изменение статических характеристик передачи. Эти изменения вызваны в основном двумя физическими процессами. Во-первых, с увеличением температуры в рабочем диапазоне температур уменьшается подвижность носителей заряда, что приводит к уменьшению тока стока. Во-вторых, происходит перераспределение носителей по энергиям и смещение уровня Ферми к середине запрещенной зоны (см. § 1.8). В связи с таким смещением уровня Ферми инверсный слой образуется у поверхности полупроводника при меньших иапряженностях электрического поля (см. рис. 1.21). Поэтому с увеличением температуры пороговое напряжение $U_{\text{ЗИ пор}}$ уменьшается. В результате статические характеристики передачи для неизменного

напряжеиия на стоке, но для разных температур пересекаются (рис. 6.11).

Таким образом, температурные изменения тока стока при неизменных напряжениях на МДП-транзисторе могут быть как отрицательными, так и положительными, а также нулевыми в определенной рабочей точке статических характеристик. Обычно эффект температурной компенсации получается при напряжениях на затворе, незначительно превышающих пороговое напряжение $U_{\text{ЗИпор}}$. Кроме того, еще надо учитывать, что крутизна характеристики S, определяющая усилительные свойства МДП-траизистора, изменяется с температурой даже при иеизменном постоянном токе стока.

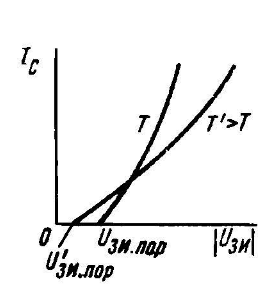


Рис. 6.11. Изменение статических характеристик передачи и порогового напряжения $U_{\text{ЗИпор}}$ МДП-транзистора с изменением температуры

МДП-траизисторы со встроенным каналом

Проводящий канал под затвором МДП-транзистора может быть создан в результате локальной диффузии или ионной имплантации соответствующих примесей в приповерхностный слой подложки. Он может возникнуть из-за перераспределения примесей вблизи поверхности полупроводниковой подложки в процессе термического окисления ее поверхности. Наконец, проводящий канал может появиться под затвором из-за фиксированного заряда в подзатворном слое диоксида кремиия, на поверхностных энергетических уровнях, а также из-за контактной разности потенциалов между металлом затвора и полупроводником подложки.

Модуляция сопротивления проводящего канала МДП-транзистора может происходить при изменении напряжения на затворе как положительной, так и отрицательной полярности. Таким

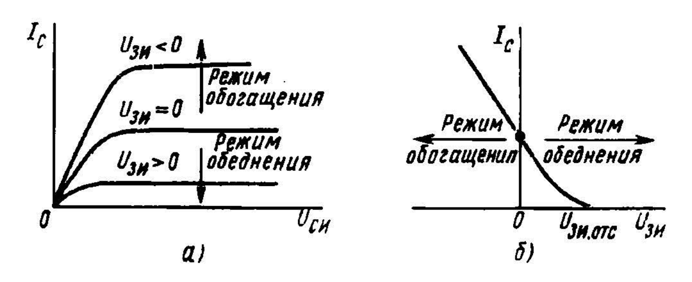


Рис. 6.12. Выходные статические характеристики (а) и статические характеристики передачи (б) МДП-транзистора со встроенным р-каналом

образом, МДП-транзистор со встроенным каналом может работать в двух режимах: *в режиме обогащения и в режиме обеднения* канала носителями заряда. Эта особенность МДП-транзисторов со встроенным каналом отражается и на смещении выходных статических характеристик при изменении напряжения на затворе и его полярности (рис. 6.12, *а*).

Статические характеристики передачи (рис. 6.12, *б*) выходят из точки на оси абсцисс, соответствующей *напряжению отсечки* $U_{\text{ЗИ отс}}$, т. е. напряжению между затвором и истоком МДП-транзистора со встроенным каналом, работающего в режиме обеднения, при котором ток стока достигает заданного низкого значения.

§ 6.6. РАСЧЕТ ВЫХОДНЫХ СТАТИЧЕСКИХ ХАРАКТЕРИСТИК ПОЛЕВОГО ТРАНЗИСТОРА С ИЗОЛИРОВАННЫМ ЗАТВОРОМ

Рассмотрим вначале удельную проводимость канала γ_s, т. е. проводимость квадрата канала вдоль оси x единичной длины ($x=1$) и единичной ширины ($y=1$) (рис. 6.13). Удельная проводимость канала является величиной, обратной удельному сопротивлению слоя (см. § 1.15):

$$\gamma_s=\frac{1}{\rho_s}=q\int_0^w p\mu_p dz=qp_s\mu_{ps}, \quad (6.17)$$

где $p_s=\int_0^w p dz$ — количество дырок в единичном квадрате канала, которое можно назвать также поверхностной плотностью дырок в канале, так как p_s имеет размерность см$^{-2}$; μ_{ps} — эффективная подвижность дырок в канале, которая из-за рассеяния на поверхности полупроводника обычно в несколько раз меньше подвижности дырок в объеме; w — толщина канала.

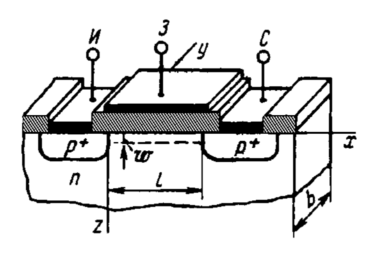


Рис. 6.13. Структура части МДП-транзистора, принятая для расчета выходных статических характеристик

Удельную проводимость канала можно представить еще в таком виде:

$$\gamma_s=qp_s\mu_{ps}=Q_{ps}\mu_{ps}, \quad (6.18)$$

где Q_{ps} — плотность подвижного заряда (дырок) в канале.

Кроме дырок в канале есть неподвижные заряды — ионизированные примеси в канале, есть неподвижные заряды в слое диоксида кремния вблизи границы раздела диоксид кремния — кремний. Таким образом, плотность полного заряда, изменяющаяся по длине канала из-за изменения потенциала канала при

прохождении тока стока, может быть определена следующим образом:

$$Q_{\text{полн}} = Q_{\text{неподв}} + Q_{ps} = C_{\text{зк}}(U_{\text{ЗИ}} - U), \qquad (6.19)$$

где U — изменяющийся по длине канала потенциал или разность потенциалов между некоторой точкой канала с координатой x и истоком; $C_{\text{зк}}$ — удельная емкость между затвором и каналом.

При напряжениях на затворе, не превышающих порогового $U_{\text{ЗИ пор}}$, канала под затвором еще нет, т. е. под затвором находится обедненный основными носителями слой, содержащий только неподвижные заряды. Поэтому

$$Q_{\text{неподв}} = C_{\text{зк}} U_{\text{ЗИ пор}}. \qquad (6.20)$$

Из соотношений (6.19) и (6.20) находим плотность подвижного заряда:

$$Q_{ps} = C_{\text{зк}}(U_{\text{ЗИ}} - U_{\text{ЗИ пор}} - U). \qquad (6.21)$$

Ток, проходящий по каналу от истока к стоку и одинаковый во всех сечениях по длине канала, или ток стока,

$$I_{\text{С}} = \gamma_s b E_x = \gamma_s b\,|dU/dx|, \qquad (6.22)$$

где b — ширина канала; E_x — изменяющаяся по длине канала продольная составляющая напряженности электрического поля.

Отсюда, используя (6.18) и (6.21), получим

$$I_{\text{С}} \int_0^l dx = \mu_{ps} C_{\text{зк}} b \int_0^{U_{\text{СИ}}} (U_{\text{ЗИ}} - U_{\text{ЗИ пор}} - U)\, dU.$$

Тогда для крутой части выходных статических характеристик

$$I_{\text{С}} = \frac{\mu_{ps} C_{\text{зк}} b}{l} \left[(U_{\text{ЗИ}} - U_{\text{ЗИ пор}})\, U_{\text{СИ}} - \frac{U_{\text{СИ}}^2}{2} \right]. \qquad (6.23)$$

Для получения выражений, соответствующих пологим частям выходных статических характеристик, надо учесть, что при потенциале стокового конца канала $U(l) \geqslant U_{\text{СИ нас}}$ происходит перекрытие канала около стока. При этом дырок (подвижного заряда) в перекрытой части канала почти нет. Они проходят перекрытую часть канала под действием большой продольной составляющей напряженности электрического поля со скоростью, равной скорости насыщения. Поэтому можно считать, что в канале около стока $Q_{ps} \approx 0$. Тогда, используя (6.21), получим

$$Q_{ps} = C_{\text{зк}}(U_{\text{ЗИ}} - U_{\text{ЗИ пор}} - U_{\text{СИ нас}}) \approx 0,$$

отсюда

$$U_{\text{СИ нас}} = U_{\text{ЗИ}} - U_{\text{ЗИ пор}}. \qquad (6.24)$$

Для пологой части выходных статических характеристик, подставляя (6.24) в (6.23), получим

$$I_С = \frac{\mu_{ps} C_{зк} b}{2l} (U_{ЗИ} - U_{ЗИ\ пор})^2. \qquad (6.25)$$

Таким образом, ток стока в первом приближеиии не зависит от напряжения иа стоке при $U_{СИ} > U_{СИ\ нас}$.

§ 6.7. ПАРАМЕТРЫ И СВОЙСТВА ПОЛЕВЫХ ТРАНЗИСТОРОВ С ИЗОЛИРОВАННЫМ ЗАТВОРОМ

Основным параметром полевого транзистора с изолированиым затвором, отражающим его усилительные свойства, является крутизиа характеристики (см. § 6.1). Крутизна характеристики передачи при иизкой частоте, соответствующая крутой части выходных статических характеристик, может быть определена путем дифференцирования (6.23) по иапряжению на затворе при неизмеииом напряжении на стоке:

$$S = \left.\frac{dI_С}{dU_{ЗИ}}\right|_{U_{СИ}=\mathrm{const}} = \frac{\mu_{ps} C_{зк} b}{l} U_{СИ}. \qquad (6.26)$$

Для пологой части выходиых статических характеристик крутизна характеристики передачи может быть получеиа путем диффереицирования (6.25) по напряжению на затворе:

$$S = \frac{\mu_{ps} C_{зк} b}{l} (U_{ЗИ} - U_{ЗИ\ нор}). \qquad (6.27)$$

Как видно, для увеличения крутизиы характеристики исходный полупроводник должеи обладать большей подвижиостью носителей заряда. Транзистор с n-каналом имеет большее значение крутизны характеристики по сравнеиию с транзистором с p-каналом, так как подвижность электронов превышает обычно подвижность дырок.

Крутизна характеристики будет больше в полевых транзисторах с меньшей длииой каиала. Нижний предел длины каиала ограничен техиологией изготовления. Обычно для изготовления полевых транзисторов с изолированным затвором применяют планарную технологию и метод фотолитографии, разрешающая способность которого ие позволяет получать длину канала меньше 3—4 мкм.

Крутизну характеристики можно увеличить путем увеличеиия удельной емкости между затвором и каналом. Эта емкость определяется относительной диэлектрической проницаемостью и толщиной слоя диэлектрика под затвором. Использование диэлектрика с большей относительной диэлектрической проницаемостью приведет к увеличению крутизны характеристики, но одновременно увеличатся и паразитные емкости между затвором и стоком, между затвором и истоком, что отрицательио повлияет

на частотные свойства полевого транзистора. Уменьшение толщины слоя диэлектрика под затвором может также привести к недопустимому уменьшению пробивного напряжения этого слоя между затвором и стоком.

Увеличение ширины канала приводит к увеличению крутизны характеристики, но одновременно и к ухудшению частотных свойств полевого транзистора в связи с увеличением паразитных емкостей.

Физическая эквивалентная схема полевого транзистора с изолированным затвором аналогична физической эквивалентной схеме полевого транзистора с управляющим переходом (см. рис. 6.5). Однако в связи с тем что затвор изолирован от полупроводника слоем диэлектрика, активные сопротивления между затвором и истоком, между затвором и стоком оказываются очень большими. Поэтому ими можно пренебречь даже на относительно малых частотах по сравнению с параллельно включенными емкостными сопротивлениями. Можно пренебречь также очень малыми сопротивлениями $r_и$ и $r_с$, которые представляют собой дифференциальные сопротивления сильнолегированных областей полупроводника под истоком и стоком.

Быстродействие полевых транзисторов с изолированным затвором определяется временем перезаряда распределенной емкости между затвором и каналом. Постоянные времени процесса перезаряда этой емкости при малом внешнем сопротивлении в цепи затвора ограничивают рабочий диапазон частот полевого транзистора с изолированным затвором частотами около 10 ГГц, т. е. принципиально такие транзисторы могут работать приблизительно до тех же частот, что и биполярные транзисторы. Основной особенностью полевых транзисторов является очень большое входное сопротивление. Активная составляющая этого сопротивления может достигать 10^{15} Ом. Поэтому полевые транзисторы применяют в схемах, имеющих также большие сопротивления.

§ 6.8. ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ С ЗАРЯДОВОЙ СВЯЗЬЮ

Прибор с зарядовой связью (ПЗС) — это полупроводниковый прибор, имеющий большое число близкорасположенных и изолированных от подложки затворов (МДП-структур), под которыми может происходить перенос к стоку информационных пакетов неосновных носителей заряда, либо инжектированных из истока, либо возникших в подложке из-за поглощения оптического излучения.

Структура и принцип действия приборов с зарядовой связью

Рассмотрим принцип действия ПЗС на примере трехтактной схемы сдвигового регистра, которую можно представить как структуру МДП-транзистора со многими затворами (рис. 6.14, *а*). Этот прибор состоит из трех секций.

Первая — *входная секция* — включает в себя исток с p^+-областью под ним и входной затвор, выполняющий роль ключа для управления движением дырок из диффузионной p^+-области истока в первую потенциальную яму.

Вторая — *секция переноса* — состоит из ряда затворов, управляющих потенциалом на границе кремний — диоксид кремния. Эти затворы соединены между собой через два. Напряжения на затворах секции переноса имеют вид импульсов различной амплитуды, которые сменяют друг друга циклической перестановкой (рис. 6.14, *б*...*д*). При таком изменении напряжения на затво-

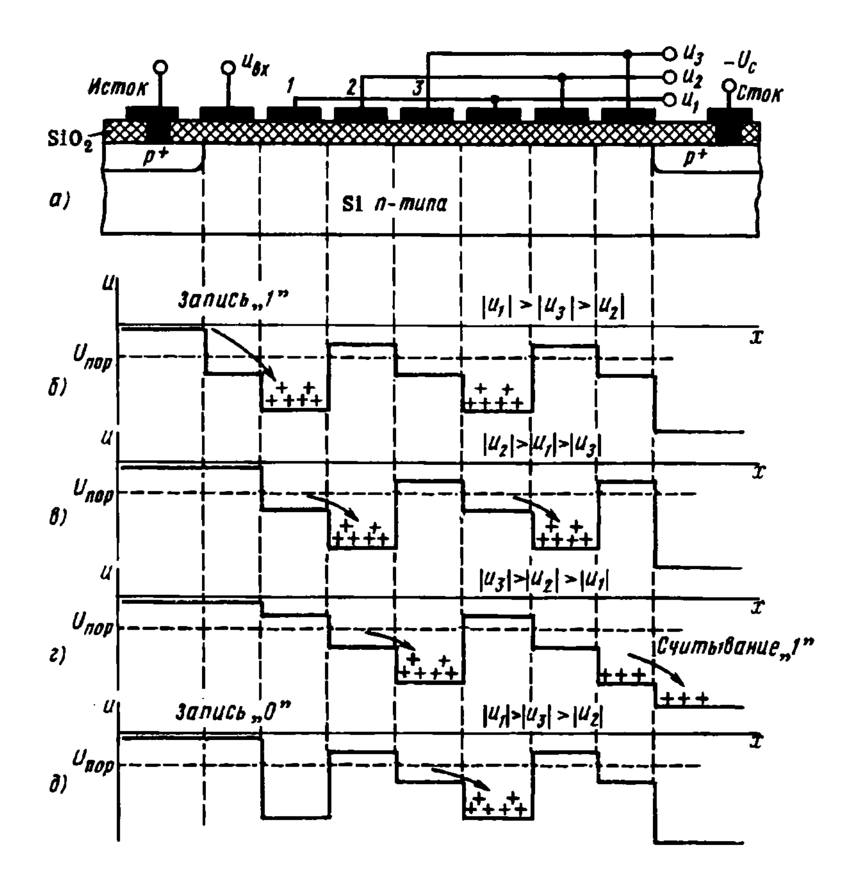


Рис. 6.14. Структура ПЗС с трехтактным питанием затворов секции переноса (*а*) и пояснение принципа его действия (*б, в, г, д*):

б — запись логической единицы путем инжекции пакета дырок в потенциальную яму под первым затвором секции переноса; *в* — перенос информационного заряда — пакета дырок в последующие потенциальные ямы при изменении потенциалов на электродах затвора; *г* — считывание логической единицы на выходе прибора при экстракции дырок из потенциальной ямы в p^+-область стока; *д* — запись логического нуля при отсутствии отрицательного потенциала на входном затворе

рах потенциальные ямы перемещаются к выходу прибора, увлекая за собой пакеты носителей заряда — дырок.

Третья — *выходная секция* — включает в себя *p-n*-переход стока. Этот переход смещен в обратном направлении и предназначен для экстракции дырок из подходящих к нему потенциальных ям (рис. 6.14, *г*).

Пусть на начальном такте работы на входной затвор подано напряжение $U_{вх}$, достаточное для образования проводящего канала под входным затвором ($|U_{вх}| > |U_{пор}|$). Если при этом на первом затворе секции переноса существует достаточно большое отрицательное напряжение, т. е. если под первым затвором секции переноса существует глубокая потенциальная яма для дырок, то дырки будут выходить из истока, проходить по каналу под входным затвором и накапливаться в потенциальной яме под первым затвором секции переноса (рис. 6.14, *б*).

Напряжение на входном затворе $U_{вх}$ снимается к началу следующего такта изменения напряжений на затворах секции переноса. Поэтому проводящий канал под входным затвором исчезает. Так происходит запись информации (например, логической единицы), которой соответствует некоторый заряд дырок $Q_{«1»}$, накопленных в потенциальной яме под первым затвором в результате инжекции из истока. Отметим, что для записи информации, соответствующей логическому нулю, на входной затвор не должно быть подано отрицательное напряжение. В этом случае не будет инжекции дырок из p^+-области истока в потенциальную яму под первым затвором (рис. 6.14, *д*) и в ней может оказаться только относительно небольшой заряд дырок $Q_{«0»}$, связанный либо с тепловой генерацией носителей заряда, либо с неполным опустошением потенциальной ямы на предыдущих тактах работы прибора.

После смены напряжений на затворах секции переноса самое отрицательное напряжение будет на втором затворе, поэтому пакет дырок передвинется в потенциальную яму под вторым затвором секции переноса (рис. 6.14, *в*). При следующих тактах изменения напряжения на затворах секции переноса будет происходить дальнейшее продвижение пакета дырок к выходной секции (рис. 6.14, *г, д*).

Если в потенциальных ямах, подходящих к *p-n*-переходу стока, нет носителей заряда — дырок, то не будет и изменения тока в цепи стока. И только в том случае, когда потенциальная яма, содержащая дырки, подойдет к *p-n*-переходу стока, произойдет экстракция этих дырок и в цепи стока пройдет импульс тока или изменится напряжение на стоке (рис. 6.14, *г*).

Параметры приборов с зарядовой связью

Следует отметить, что ПЗС является типично динамическим устройством и имеет нижний и верхний пределы тактовых частот импульсов напряжения, питающих секцию переноса.

Нижний предел тактовой частоты определяется тем, что между потенциальной ямой у поверхности и остальным объемом полупроводника проходят токи, связанные с тепловой генерацией носителей заряда и ничем в принципе не отличающиеся от обратного тока экстракции через *p-n*-переход. Эти токи влияют на уровень логического нуля, повышая заряд дырок в пустых потенциальных ямах. В зависимости от температуры и свойств полупроводника заметное накопление дырок в пустых потенциальных ямах может произойти за время от сотых долей до единиц секунд. Поэтому нижний предел тактовой частоты ПЗС составляет обычно единицы — десятки килогерц.

Верхний предел тактовой частоты определяется временем перетекания заряда из одной потенциальной ямы в другую (порядка единиц наносекунд). За более короткое время весь заряд не успевает перейти из одной потенциальной ямы в соседнюю. Поэтому верхний предел тактовых частот для ПЗС определяется обычно десятками мегагерц.

В диапазоне рабочих частот, далеких от предельных, в ПЗС тоже не происходит полной передачи информационного заряда из одной потенциальной ямы в другую. Связано это с явлениями захвата носителей заряда поверхностными энергетическими уровнями ловушек захвата. В результате, во-первых, уменьшается общий заряд передаваемого пакета дырок, т. е. уменьшается уровень логической единицы. Во-вторых, захваченные ловушками захвата и освободившиеся через некоторое время носители заряда могут попасть в пустые потенциальные ямы, исказив тем самым уровень логического нуля. Для уменьшения влияния этого эффекта необходимо обеспечить плотность поверхностных состояний примерно на два порядка меньшую, чем допускается при производстве МДП-транзисторов.

Другим, более простым, методом борьбы с рассмотренным эффектом является кодирование логического нуля не отсутствием заряда в потенциальной яме, а некоторым малым зарядом. Этот заряд во время продвижения по секции переноса ПЗС приводит к заполнению большей части поверхностных состояний, которые уже не могут больше захватывать носители заряда при прохождении пакета дырок. При этом, однако, уменьшается амплитуда сигнала на выходе ПЗС.

Третьим методом борьбы с явлением захвата является использование так называемого углубленного канала. При этом путем введения соответствующих примесей в поверхностный слой полупроводника создают такое распределение электрического поля вблизи поверхности, при котором носители заряда движутся не

вблизи поверхности, а на некоторой глубине (до нескольких микрометров). Это резко снижает потери на захват носителей. К тому же в ПЗС с углубленным каналом удается сдвинуть верхний предел рабочих частот по крайней мере за 100 МГц. Однако приборы с углубленным каналом имеют значительно меньшую эффективность управления со стороны затворов и соответственно меньший информационный заряд по сравнению с обычными ПЗС.

Для оценки рассмотренного эффекта захвата носителей в ПЗС используют параметр *эффективности передачи заряда* $\eta=(Q_{«1»}-Q_{«0»})_{i+1}/(Q_{«1»}-Q_{«0»})_i$, показывающий, какая доля заряда переносится из одной потенциальной ямы i в другую $i+1$. Значение эффективности передачи η обычно близко к 1, поэтому удобнее пользоваться *коэффициентом потерь* (неэффективностью передачи) $K_п=1-\eta$.

Для хороших ПЗС коэффициент потерь составляет менее 10^{-4}. Частотная зависимость коэффициента потерь показана на рис. 6.15. Спад коэффициента потерь $K_п$ в диапазоне низких частот определяется уменьшением влияния обратных токов между потенциальной ямой и остальным объемом полупроводника (подложкой). Подъем коэффициента потерь при высоких частотах связан с неполным переносом заряда из одной потенциальной ямы в другую из-за быстрого изменения тактовых импульсов на затворах. Конечное значение коэффициента потерь в диапазоне средних частот определяется эффектом захвата носителей заряда ловушками захвата.

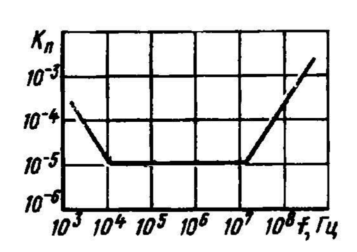


Рис. 6.15. Частотная зависимость коэффициента потерь ПЗС

Как видно, в ПЗС неизбежно происходит потеря информационного заряда. Чтобы устранить этот недостаток, используют схемы регенерации, которые представляют собой, по существу, усилители. Считываемый с ПЗС сигнал усиливается с соответствующим формированием его уровней, а затем производится запись информации в цепочку ПЗС. Для длительного хранения информации цепочки ПЗС замыкают в кольцо. Регенерация информационного заряда может быть совмещена с выводом информации — тогда получается прибор с неразрушающим считыванием информации.

Через управляющие электроды затворов ПЗС практически не проходит ток электропроводности, так как затворы изолированы от полупроводника (подложки) слоем хорошего диэлектрика — диоксида кремния. Но в цепи электродов секции переноса все-таки расходуется некоторая мощность на перенос информационного заряда. Эта мощность пропорциональна тактовой частоте.

С точки зрения конструкции и технологии ПЗС отличаются от

других приборов тем, что имеют малое число диффузионных областей и контактов металла с полупроводником, т. е. малое число потенциальио неиадежиых элемеитов структуры. Это приводит к высокому процеиту выхода годиых приборов при изготовлеиии ПЗС, к низкой их себестоимости и большой надежности.

Применение приборов с зарядовой связью

К настоящему времеии выявилось три осиовиых направлеиия использования ПЗС: 1) запомииающие устройства электроиио-вычислительиых машии; 2) устройства преобразоваиия изображеиия в электрические сигиалы; 3) устройства обработки аиалоговой иииформации.

Запоминающие устройства ЭВМ. По принципу действия ПЗС представляют собой запоминающие устройства типа линий задержки. Запомииающие устройства иа базе ПЗС доведены до промышлениого выпуска, так как оии лучше всего соответствуют природе ПЗС — сдвигающие регистры с последовательиым вводом и выводом ииформации.

Информация иепрерывио циркулирует в таком запомииающем устройстве на ПЗС с регенерацией. При обращеиии к запоминающему устройству производится выборка записаиной ииформации с регенерацией или без иее, т. е. с иеразрушающим считыванием или с разрушеиием записаиной ииформации. Часто в режиме храиеиия ииформация циркулирует по прибору отиосительио медлеино при тактовой частоте порядка десятка килогерц, чтобы только обеспечить регеиерацию и ие тратить большой мощиости иа перенос ииформациониого заряда. При обращеиии к запоминающему устройству тактовая частота повышается до предельной — около нескольких мегагерц, что обеспечивает быструю выборку записаиной ииформации.

Устройства преобразоваиия изображения в электрические сигналы. Прииции действия таких устройств основаи иа том, что при освещеиии ПЗС в полупроводнике около его поверхиости образуются пары иосителей заряда электрон — дырка, которые разделяются электрическим полем потенциальиой ямы под затвором секции переиоса. Образующиеся при поглощении кваитов света носители заполияют потенциальные ямы пропорциоиальио освещеииости даииой области ПЗС. Если затем произвести обычиым путем сдвиг записаниой световой информации, то сигиал иа выходе ПЗС будет повторять распределеиие освещеииости, т. е. будет выделеиа строка изображеиия. Так же может быть выделена следующая строка и т. д. В настоящее время создаиы передающие камеры с ПЗС, достигающие обычиого телевизиоиного стандарта по разрешающей способиости, в том числе и для цветного телевидения.

Устройства обработки аиалоговой информации. С помощью ПЗС могут запомииаться и аналоговые сигналы, ио в этом случае

становится невозможной регенерация записанной информации. Однако и простое запоминание открывает большие возможности использования ПЗС, так как эти приборы позволяют регулировать задержку переноса информации. Простейшим вариантом использования ПЗС для обработки аналоговой информации оказались линии фиксированной задержки для телевизионных приемников цветного изображения.

§ 6.9. РАЗНОВИДНОСТИ ПРИБОРОВ С ЗАРЯДОВОЙ СВЯЗЬЮ

Принцип действия приборов с зарядовой связью был рассмотрен в § 6.8 на примере ПЗС с трехтактным питанием затворов секции переноса. Конструкция такого прибора имеет ряд недостатков:

1) металлические электроды — затворы — необходимо располагать на очень малом расстоянии друг от друга (около 2—3 мкм), что затрудняет их изготовление;

2) при малом расстоянии между затворами возможно загрязнение тонкого слоя диоксида кремния, расположенного между электродами, примесями из атмосферы;

3) неустранимые пересечения линий металлизации для соединения затворов между собой усложняют технологию.

Эти недостатки частично можно устранить в других конструкциях подобных приборов.

Двухтактные приборы с зарядовой связью

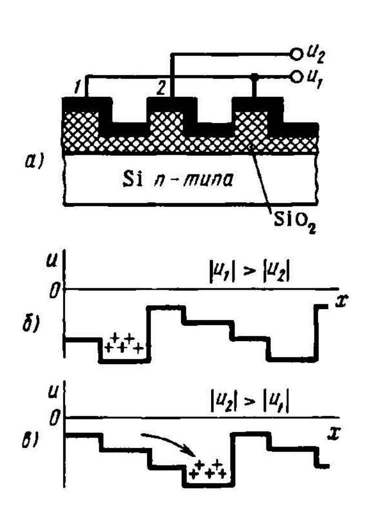


Рис. 6.16. Структура секции переноса ПЗС со ступенчатым диэлектриком (*а*) и пояснение принципа переноса информационного заряда при двухтактном питании затворов секции переноса (*б, в*)

Структура двухтактного ПЗС или ПЗС со ступенчатым диэлектриком отличается чередованием областей полупроводника (кремния) с толстым и тонким слоями диоксида кремния (рис. 6.16, *а*). Электрод каждого затвора секции переноса расположен на поверхности толстого и тонкого слоев диоксида кремния. При приложении напряжения на затвор под ним автоматически получается асимметричная потенциальная яма, конфигурация которой обеспечивает направленное перемещение пакета носителей заряда к стоку (рис. 6.16, *б, в*). Понятно, что схема управления секцией переноса такого ПЗС значительно проще, чем у трехтактного ПЗС. Соединение затворов между собой в двухтактных ПЗС может быть выполнено

без пересечений лииий металлизации, что упрощает технологию изготовления таких приборов.

ПЗС с рядом скрытых затворов (с двухслойной металлизацией)

В структуре секции переиоса такого прнбора часть затворов сделана из поликристаллического кремния (или из тугоплавкого металла, иапример молибдена). В зазорах между кремниевыми затворами на поверхиости слоя диоксида кремния расположены металлические затворы (рис. 6.17), которые отделены от затворов из поликристаллического кремния слоем диоксида кремния, наиесенным после создания затворов из поликристаллического кремиия. В такой структуре удается сократить расстояние между электродами затворов до зиачеиия, равного толщине слоя диоксида кремиия.

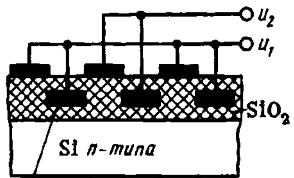


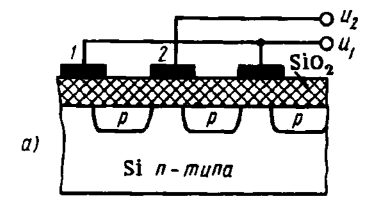
Рис. 6.17. Структура секции переноса ПЗС с рядом скрытых затворов из поликристаллического кремиия

ПЗС на цепочках МДП-транзисторов

Структура секции переноса таких ПЗС представляет собой цепочку МДП-транзисторов с двухтактиыми цепями управлеиия (рис. 6.18, *а*). Диффузионные области с электропроводиостью *р*-типа, являющиеся стоком для одного и истоком для другого МДП-транзистора, не подключаются к источнику питания в отличие от обычных МДП-транзисторов.

Информациониый заряд в виде пакета дырок может храииться в диффузионных областях с электропроводностью *р*-типа, так как потенциальный барьер *р-п*-перехода между диффузионной областью и подложкой препятствует растекаиию дырок в режиме хранения, несмотря на некоторое прямое смещение этого перехода из-за положительного заряда пакета дырок в диффузионной *р*-области (рис. 6.18, *б*).

В режиме передачи информации (пакета дырок) на каждый второй затвор поступает импульс отрицательной полярности, амплитуда которого выше значения $U_{пор}$. По этой причине под каждым вторым затво-



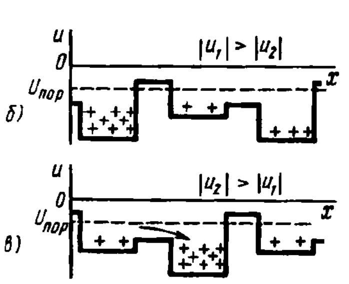


Рис. 6.18. Структура секции переноса ПЗС на цепочках МДП-транзисторов (*а*) и распределение поверхностиого потенциала в режиме переноса информационного заряда (*б, в*)

ром в полупроводнике возникает канал. Из-за несимметрии расположения металлического электрода относительно диффузионных областей каждая вторая диффузионная область оказывается под бо́льшим отрицательным потенциалом, что обеспечивает направленное движение дырок из одной диффузионной области в следующую, т. е. в более глубокую потенциальную яму (рис. 6.18, *в*).

Таким образом, принцип действия ПЗС на цепочках МДП-транзисторов аналогичен принципу действия ПЗС других конструкций и структур, которые были рассмотрены ранее. Отличие ПЗС на цепочках МДП-транзисторов от других ПЗС состоит в том, что структура их секций переноса содержит достаточно большое число *p*-*n*-переходов, что делает ее менее совершенной с технологической точки зрения.

Поверхностно-зарядовый транзистор

Простейшая структура поверхностно-зарядового транзистора состоит из кремниевой подложки и трех изолированных друг от друга и от подложки электродов (рис. 6.19). Первый и третий электроды могут быть скрытыми. Их делают обычно из поликристаллического кремиия или тугоплавкого металла (например, молибдена). При подаче на первый или третий электроды отрицательного потенциала в полупроводнике под соответствующим электродом образуется потенциальная яма для неосновных носителей — дырок. Информация в виде пакета дырок может быть введена в эту потенциальную яму путем освещения. Таким образом, первый и третий электроды вместе с областями полупроводника под этими электродами могут выполнять функции приема и хранения информационного заряда.

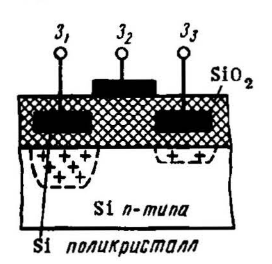


Рис. 6.19. Структура поверхностно-зарядового транзистора при хранении информационного заряда под первым электродом

Второй электрод выполняет функцию затвора — при подаче на него отрицательного потенциала образуется канал, соединяющий потенциальные ямы под первым и третьим электродами. При этом может происходить перетекание информационного заряда из потенциальной ямы, например, под первым электродом в потенциальную яму под третьим электродом, если потенциал третьего электрода более отрицателен.

Так же как в ПЗС других конструкций, при статическом режиме работы поверхностно-зарядового транзистора происходит накопление паразитного заряда в пустой потенциальной яме из-за тепловой генерации носителей заряда, т. е. в статическом режиме такой прибор работать не может.

Контрольные вопросы

1. Какие разновидности полевых транзисторов существуют?

2. Как с помощью полевого транзистора происходит преобразование энергии относительно мощного источника питания выходной цепи в энергию электрических колебаний?

3. Почему свойства и характеристики полевых транзисторов следует описывать системой уравнений, в которых токи являются функциями напряжений, а не наоборот?

4. Какие физические факторы могут влиять на характер зависимости тока стока от напряжения на стоке полевого транзистора с управляющим переходом?

5. Какими физическими явлениями, происходящими в полевом транзисторе, ограничивается диапазон рабочих частот этого прибора?

6. Чем отличаются структуры МДП-транзисторов с индуцированным и со встроенным каналами? Как это отличие отражается на статических характеристиках передачи и каковы специфические параметры тех и других полевых транзисторов?

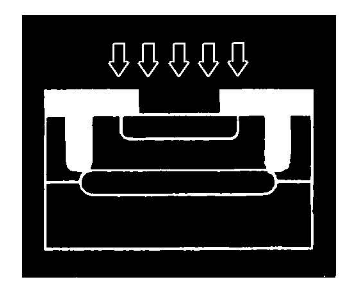
7. Каков принцип действия приборов с зарядовой связью?

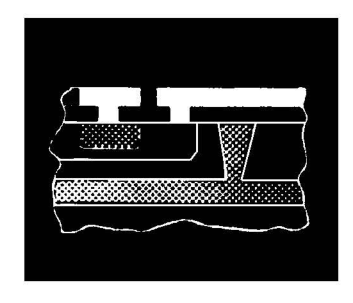
8. Какие существуют разновидности структур секции переноса приборов с зарядовой связью?

9. Каков смысл основных параметров прибора с зарядовой связью?

10. Какие факторы влияют на эффективность переноса информационного заряда в приборе с зарядовой связью? Как и почему коэффициент потерь зависит от частоты тактовых импульсов, питающих секцию переноса прибора с зарядовой связью?

11. Каковы основные применения приборов с зарядовой связью?





Интегральные микросхемы

§ 7.1. ЗАДАЧИ И ПРИНЦИПЫ МИКРОЭЛЕКТРОНИКИ

Микроэлектроника — это научно-техническое направление электроники, охватывающее проблемы исследования, конструирования и изготовления высоконадежных и экономичных микроминиатюрных электронных схем и устройств с помощью комплекса физических, химических, схемотехнических и других методов.

Первой задачей микроэлектроники является создание максимально надежных электронных схем и устройств. Эта задача решается в основном путем использования качественно новых принципов изготовления электронной аппаратуры, т. е. путем отказа от использования дискретных элементов электронной аппаратуры и создания интегральных микросхем, в которых формирование активных элементов (транзисторов, диодов), пассивных элементов (резисторов, конденсаторов) и соединительных элементов электронной схемы происходит на поверхности или в объеме полупроводникового кристалла или на поверхности диэлектрической подложки в едином технологическом цикле. Минимальное количество внутрисхемных соединений дает возможность резко повысить надежность микроэлектрониой аппаратуры. Именно этим преодолеваются сложные противоречия между возросшими требованиями к надежности электрониой аппаратуры и ее стремительным усложнением.

Второй задачей микроэлектроники является снижеиие стоимости электронных схем и устройств. Эта задача решается путем формироваиия за единый технологический цикл структур различных элементов, межэлементных соединений и контактных площадок для многих интегральных микросхем на относительно большой полупроводииковой пластине или на диэлектрической подложке с последующим разделением соответствеиио на кристаллы или на платы интегральных микросхем. При этом удается исклю-

чить много нерациональных технологических операций, сократить число внутрисхемных соединений, исключить раздельную герметизацию отдельных элементов и существенно сократить число сборочных операций, которые необходимы при изготовлении дискретных элементов и при их сборке в схему. Эти преимущества интегральных микросхем становятся более значительными по мере их усложнения и возрастания в них числа элементов.

Наряду с решением этих двух важнейших задач микроэлектроники создание и использование интегральных микросхем приводит к резкому уменьшению массы и объема электронной аппаратуры по сравнению с массой и объемом аппаратуры на дискретных элементах, а также к уменьшению потребляемой мощности. Дальнейшее уменьшение массы и объема интегральных микросхем рассматривается как задача второстепенной важности.

Интегральная микросхема (микросхема) — это микроэлектронное изделие, выполняющее определенную функцию преобразования, обработки сигнала и (или) накапливания информации и имеющее высокую плотность упаковки электрически соединенных элементов, которое с точки зрения требований к испытаниям, приемке, поставке и эксплуатации рассматривается как единое целое.

Плотность упаковки интегральной микросхемы — это отношение числа элементов интегральной микросхемы к объему интегральной микросхемы без учета объема выводов. Другим параметром, характеризующим степень сложности интегральной микросхемы или число содержащихся в ней элементов, является *степень интеграции*. Если интегральная микросхема содержит до 10 элементов включительно, то ее называют интегральной микросхемой первой степени интеграции; свыше 10 до 100 элементов — второй степени интеграции; свыше 100 до 1000 элементов — третьей степени интеграции и т. д. Интегральную микросхему, содержащую 500 элементов и более, изготовленных по биполярной технологии, или 1000 элементов и более, изготовленных по МДП-технологии, называют *большой интегральной микросхемой* (БИС).

§ 7.2. КЛАССИФИКАЦИЯ ИНТЕГРАЛЬНЫХ МИКРОСХЕМ

Интегральные микросхемы классифицируют по технологии изготовления, по функциональному назначению и по другим признакам.

По конструктивно-технологическому признаку различают полупроводниковые и гибридные интегральные микросхемы.

Полупроводниковая интегральная микросхема — это интегральная микросхема, все элементы и межэлементные соединения которой выполнены в объеме и на поверхности полупроводника.

Основными активными элементами полупроводниковых интегральных микросхем могут быть либо биполярные транзисторы, либо полевые транзисторы, в качестве которых обычно использу-

ют МДП-транзисторы с индуцированным каналом. Поэтому различают биполярные и МДП интегральные микросхемы. Элементы биполярной интегральной микросхемы должны быть изолированы друг от друга для исключения паразитного взаимодействия. Методы изоляции рассмотрены в § 7.3. В связи с особенностями МДП-транзисторов элементы МДП интегральных микросхем не нуждаются в специальной изоляции друг от друга.

Соединения отдельных элементов между собой, необходимые для функционирования схемы, осуществляют с помощью тонких

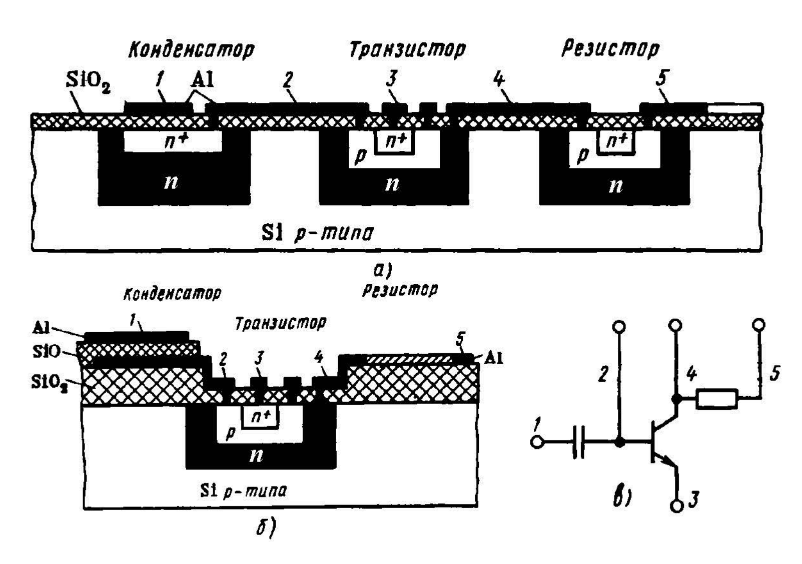


Рис. 7.1. Варианты структур полупроводниковых интегральных микросхем с различным выполнением пассивных элементов (*а*, *б*) и эквивалентная схема этих структур (*в*)

металлических полосок, нанесенных на окисленную поверхность кристалла. Примеры структур биполярных интегральных микросхем показаны на рис. 7.1. Примером МДП интегральной микросхемы может служить прибор с зарядовой связью, рассмотренный в § 6.8.

Гибридная интегральная микросхема — это интегральная микросхема, часть которой может быть выделена как самостоятельное изделие с точки зрения требований к испытаниям, приемке, поставке и эксплуатации.

Часть гибридной интегральной микросхемы, которая может быть выделена как самостоятельное изделие, называют *компонентом интегральной микросхемы* (в отличие от *элемента*, который выполнен нераздельно от кристалла полупроводниковой интегральной микросхемы или от подложки гибридной интегральной микросхемы). В состав гибридной интегральной микросхемы

могут входить в качестве компонентов не только транзисторы или диоды, но и целые полупроводниковые интегральные микросхемы. Пассивные элементы гибридных интегральных микросхем изготовляют обычно на ситалловой, керамической или стеклянной подложке путем нанесения различных диэлектрических, резистивных и металлических пленок. На этой же подложке выполняют межэлементные и межкомпонентные соединения, а также контактные площадки. Пример гибридной интегральной микросхемы показан на рис. 7.2.

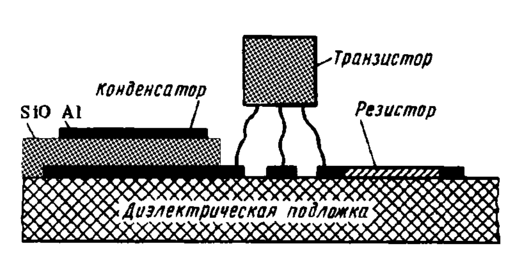


Рис. 7.2. Структура гибридной интегральной микросхемы, эквивалентная схема которой показана на рис. 7.1, *в*.

Проектирование и изготовление гибридных интегральных микросхем целесообразно для решения специальных, частных задач при относительно малом количестве требуемых изделий. Проектирование пленочных резисторов и конденсаторов на диэлектрической подложке может быть проведено оперативно, а изготовление этих элементов не требует дорогостоящего оборудования.

Если же гибридная интегральная микросхема состоит из нескольких полупроводниковых интегральных микросхем, то ее отличительной особенностью может быть, наоборот, многофункциональность. Производство таких гибридных интегральных микросхем должно быть массовым.

По функциональному назначению все интегральные микросхемы принято делить на аналоговые и цифровые.

Аналоговая интегральная микросхема (аналоговая микросхема) — это интегральная микросхема, предназначенная для преобразования и обработки сигналов, изменяющихся по закону непрерывной функции.

В основе аналоговых схем вообще и аналоговых интегральных микросхем в частности лежат простейшие усилительные каскады. Используя много каскадов, создают различные усилители, стабилизаторы напряжения и тока, преобразователи частоты, фазы, длительности, генераторы синусоидальных, прямоугольных и других сигналов, а также другие схемы.

Цифровая интегральная микросхема (цифровая микросхема) — это интегральная микросхема, предназначенная для преобразования и обработки сигналов, изменяющихся по закону дискретной функции.

В основе цифровых интегральных микросхем лежат транзисторные ключи, способные находиться в двух устойчивых состояниях: открытом и закрытом. Использование транзисторных ключей дает возможность создавать различные *логические, триггерные* и другие интегральные микросхемы. Цифровые интегральные микросхемы применяют в устройствах дискретной обработки

информации электронно-вычислительных машин (ЭВМ), системах автоматики и т. п.

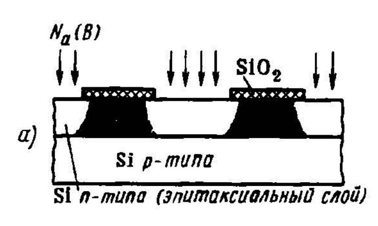
§ 7.3. МЕТОДЫ ИЗОЛЯЦИИ ЭЛЕМЕНТОВ ИНТЕГРАЛЬНЫХ МИКРОСХЕМ

Все элементы полупроводниковых интегральных микросхем и компонентов гибридных интегральных микросхем, выполненные в объеме кристалла, должны быть изолированы друг от друга для исключения паразитного взаимодействия между ними. Исключение составляют МДП-транзисторы, о чем уже было сказано. Для разделения отдельных элементов используют изоляцию их *p-n*-переходом, смещенным в обратном направлении, и иногда диэлектрическую изоляцию.

Изоляция элементов электронно-дырочным переходом

Принцип такой изоляции заключается в том, что для каждого элемента в кристалле формируется свой так называемый островок, окруженный *p-n*-переходом, который при работе микросхемы смещается в обратном направлении. Ток утечки такой изоляции обычно не превышает 10^{-7} А.

Электронно-дырочный переход, изолирующий отдельные элементы полупроводниковой интегральной микросхемы друг от друга, может быть создан различными способами, которых в настоящее время разработано более десятка. Рассмотрим самый распространенный из них — *планарно-эпитаксиальный с разделительной диффузией* (см. рис. 7.3). Вначале на пластине кремния выращивают эпитаксиальный слой с электропроводностью типа, противоположного типу электропроводности объема полупроводника, т. е. создают эпитаксиальный *p-n*-переход по всей площади пластины кремния. Затем проводят обычные этапы планарной технологии: окисление поверхности эпитаксиальной пленки, нанесение фоторезиста, засветку его через маску — фотошаблон, вскрытие окон в диоксиде кремния, локальную диффузию акцепторов через весь эпитаксиальный слой до подложки с электропроводностью *p*-типа (рис. 7.3, *а*).



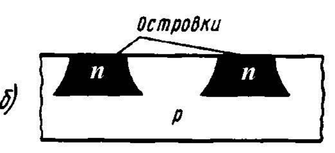


Рис. 7.3. Формирование островков с электропроводностью *n*-типа на монокристаллической подложке кремния с электропроводностью *p*-типа планарно-эпитаксиальным методом с разделительной диффузией

В полученных островках с электропроводностью *n*-типа (рис. 7.3, *б*) в дальнейшем формируют структуры транзисторов

или других полупроводниковых приборов. Для улучшения некоторых параметров и характеристик транзисторов перед выращиванием эпитаксиального слоя может быть проведена дополнительная локальная диффузия донорной примеси методом планарной технологии с целью создания скрытых легированных слоев (n^+-слоев). В результате под коллектором будущего транзистора появляется хорошо проводящая область, что и приводит к улучшению характеристик транзисторов.

Диэлектрическая изоляция

Метод диэлектрической изоляции также имеет много вариантов. Рассмотрим *метод изоляции элементов диоксидом кремния*. На рис. 7.4 показана последовательность операций при использовании этого метода: окисление монокристаллической пластины кремния (рис. 7.4, *а*); фотолитография; вытравливание канавок в кремнии через вскрытые в диоксиде кремния окна — глубина канавок около 20 мкм (рис. 7.4, *б*); повторное окисление кремния при высокой температуре или нанесение диоксида кремния другим способом — толщина слоя диоксида около 1 мкм (рис. 7.4, *в*); выращивание на слое диоксида кремния поликристаллического кремния толщиной 100...200 мкм путем, например, пиролитического разложения силана (рис. 7.4, *г*); сошлифовка или стравливание с противоположной стороны монокристаллического кремния до разделения островков (рис. 7.4, *д*).

Метод диэлектрической изоляции позволяет получить хорошую изоляцию как по постоянному, так и по переменному току, поскольку емкость, связанная с оксидным слоем, может быть очень малой (300 пФ/мм^2 при слое диоксида толщиной 1 мкм). Пробивное напряжение для диэлектрической изоляции получается значительно большим по сравнению с пробивным напряжением для изоляции *p-n*-переходом (выше 800 В).

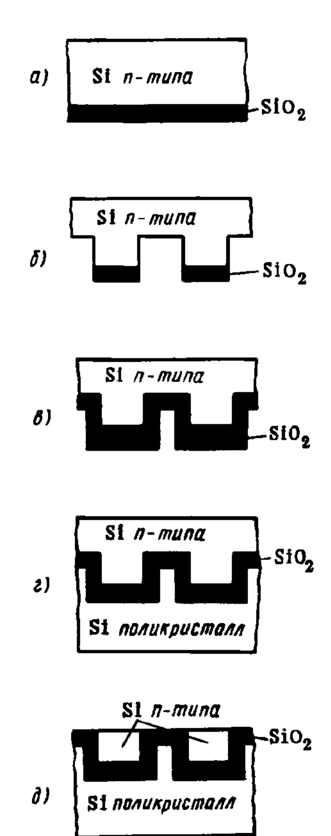


Рис. 7.4. Последовательность основных технологических этапов формирования островков монокристаллического кремния на поликристаллической подложке кремния методом диэлектрической изоляции:

а — окисление монокристаллического кремния; *б* — вытравливание канавок в кремниевой пластине через окна в слое диоксида кремния; *в* — повторное окисление кремния; *г* — наращивание поликристаллического кремния; *д* — сошлифовка монокристаллического кремния до разделения островков

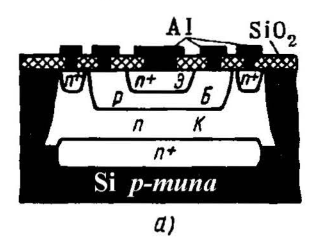
Основным недостатком этого метода изоляции является его бóльшая стоимость по сравнению с планарно-эпитаксиальным.

§ 7.4. АКТИВНЫЕ ЭЛЕМЕНТЫ

В качестве активных элементов интегральных микросхем используют обычно различные транзисторные структуры, сформированные в кристаллах кремния методами планарной технологии. Транзисторы интегральных микросхем могут отличаться не только по структуре, но и по принципу действия.

Биполярный транзистор

Биполярный транзистор является распространенным активным элементом в современных интегральных микросхемах. Структура биполярного транзистора в интегральных микросхемах (интегрального транзистора) отличается от структуры дискретного транзистора изоляцией от подложки. Другая особенность связана с тем, что вывод от коллекторной области интегрального транзистора осуществляется на верхней поверхности кристалла. Поэтому для уменьшения объемного сопротивления области коллектора перед эпитаксиальным наращиванием производится обычно подлегирование подложки в тех местах, где будут сформированы транзисторные структуры, т. е. создается скрытый n^+-слой (рис. 7.5).



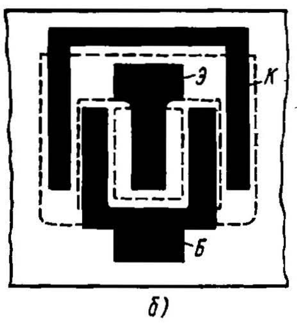


Рис. 7.5. Структура биполярного транзистора со скрытым n^+-слоем (*а*) и топология электродов этого транзистора (*б*)

Однако даже при наличии скрытого n^+-слоя сопротивление коллекторной области интегрального транзистора оказывается больше аналогичного сопротивления дискретного транзистора, так как скрытый n^+-слой отделен от коллекторного электрода высокоомным слоем коллекторной области. Это приводит к некоторому ухудшению частотных свойств интегрального транзистора в связи с увеличением постоянной времени цепи коллектора (времени перезаряда барьерной емкости коллектора). Поэтому граничные частоты $f_{гр}$ биполярных транзисторов в интегральных микросхемах обычно не превышают 500 МГц. При этом необходимо также учесть, что выходная емкость интегрального транзистора состоит не только из барьерной емкости коллекторного перехода, но и из барьерной емкости изолирующего перехода между областью коллектора интегрального транзистора и остальной частью кристалла.

Кроме того, в связи с увеличенным сопротивлением коллекторной области интегральный транзистор имеет повышенное значение напряжения между коллектором и эмиттером в режиме насыщения $U_{КЭнас}$.

Интегральный транзистор отличается от аналогичного дискретного транзистора, изготовленного по той же планарной технологии, значительно меньшими размерами. Объясняется это тем, что в дискретном планарном транзисторе должны быть контактные площадки для подсоединения проволочных выводов эмиттера, базы и коллектора. Размер контактных площадок должен быть не менее 100×100 мкм. В интегральной микросхеме контактные площадки нужны только для подсоединения выводов от всей схемы. Отдельные же элементы интегральной микросхемы соединены между собой межэлементными соединениями в виде тонких и узких (несколько микрометров) металлических полос.

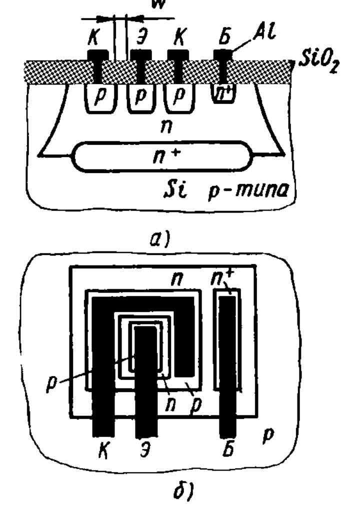


Рис. 7.6. Структура горизонтального транзистора *p-n-p*-типа (*а*) и топология этого транзистора (*б*)

Основу биполярных интегральных микросхем составляют транзисторы *n-p-n*-типа, что вызвано удобствами формирования именно *n-p-n*-структур и несколько лучшими параметрами интегральных *n-p-n*-транзисторов по сравнению с параметрами интегральных транзисторов *p-n-p*-типа.

Дело в том, что для формирования сильнолегированных эмиттерных областей транзисторов *n-p-n*-типа обычно используют диффузию фосфора, который имеет большую растворимость в кремнии и относительно малый коэффициент диффузии. Таким образом, для формирования *p-n-p*-транзистора в интегральной микросхеме, содержащей *n-p-n*-транзисторы, необходимо еще провести дополнительную диффузию какого-то акцептора с предельной растворимостью, превышающей предельную растворимость фосфора. А такие акцепторы практически отсутствуют.

Поэтому основным приемлемым вариантом интегрального транзистора *p-n-p*-типа является так называемый *горизонтальный* или *боковой транзистор* (рис. 7.6). Для его формирования не надо вводить дополнительных технологических операций, так как *p*-области его эмиттера и коллектора получаются одновременно при создании *p*-области базы транзистора *n-p-n*-типа. Однако горизонтальный *p-n-p*-транзистор оказывается бездрейфовым из-за однородного легирования его базовой области — эпитаксиального слоя. Толщина активной части базы горизонтального

траизистора получается отиосительио большой. Все это приводит к посредственным частотиым свойствам горизонтального транзистора: его граничная частота не превышает обычно нескольких десятков мегагерц.

Из структуры горизонтального транзистора (рис. 7.6) ясно, что для получения большего коэффициеита передачи тока эмиттера или базы необходимо, чтобы площадь донной части эмиттерной области была мала по сравнению с площадью боковых частей этой области. Значит, эмиттерная область должна быть по возможности более узкая (ширина окна в слое диоксида кремния для диффузии акцепторов должна составлять 3...5 мкм).

У горизонтального транзистора должны быть одинаковыми пробивные напряжения эмиттерного и коллекторного переходов. Близкими должны быть и коэффициенты передачи тока эмиттера при нормальном и инверсном включении такого транзистора, так как области эмиттера и коллектора одинаковы по свойствам.

Горизонтальная структура позволяет легко осуществить *многоколлекторный транзистор*. Для этого достаточно кольцевую область коллектора разделить на несколько частей и предусмотреть отдельные выводы от каждой части — от каждого коллектора. Коэффициент передачи тока для каждого коллектора будет, конечно, в соответствующее число раз меньше, чем для единого коллектора, но все коллекторы будут действовать «синхронно», а иагрузки во всех коллекторных цепях будут электрически разделены. Многоколлекторный транзистор оказывается удобным для некоторых цифровых интегральных микросхем.

Биполярный транзистор с диодом Шотки

Биполярный транзистор в цифровых интегральнных микросхемах обычно выполняет функцию ключа и все время работает либо в режиме насыщения, либо в режиме отсечки. В режиме насыщения происходит накопление неосновных носителей заряда в базе транзистора, а также в коллекторной области (см. § 4.15). Процессы накопления неосновных носителей и их последующего рассасывания при переводе транзистора в режим отсечки или в выключенное состояние связаны с относительно медленным процессом диффузии неосновных носителей заряда. Инерционность этих процессов определяет скорость переключеиия транзистора из включенного состояния в выключенное и обратно, т. е. скорость срабатывания схемы.

Для ускорения процесса накопления и рассасывания неосновных носителей заряда целесообразно ограничить их накопление. Достичь этого можно путем шунтирования коллекторного перехода транзистора диодом Шотки, т. е. диодом с выпрямляющим электрическим переходом между металлом и полупроводником (см. § 3.22). Структура такого интегрального транзистора и его эквивалентная схема показаны на рис. 7.7.

Алюминиевый электрод образует с p-областью базы омический переход, а переход между алюминиевым электродом и относительно высокоомной n-областью коллектора получается выпрямляющим. Из-за неравенства работ выхода электронов из алюминия и из кремния с электропроводностью n-типа и в результате химической обработки поверхности кремниевого кристалла на контакте для электронов возникает потенциальный барьер высотой около 0,6 эВ (рис. 7.8, *а*), что несколько меньше

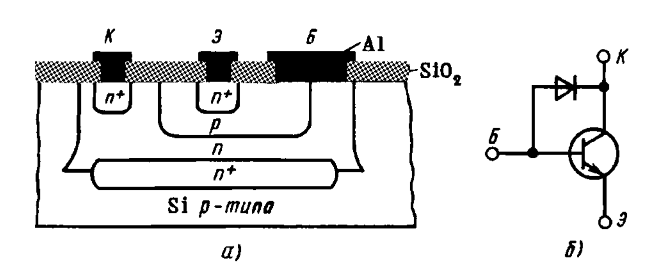


Рис. 7.7. Структура транзистора с диодом Шотки (*а*) и его эквивалентная схема (*б*)

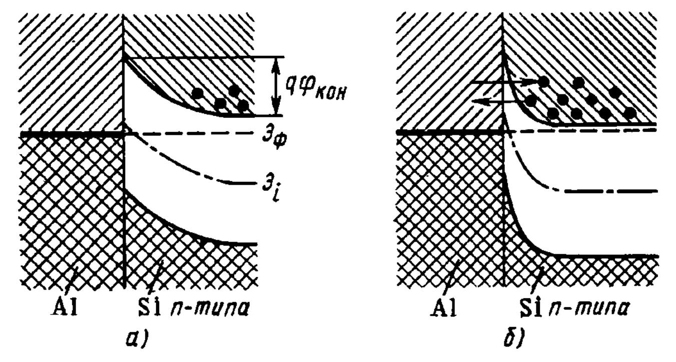


Рис. 7.8. Энергетическая диаграмма выпрямляющего перехода между алюминиевым электродом и высокоомной n-областью коллектора (*а*) и омического перехода между алюминиевым электродом и сильнолегированной n^+-областью эмиттера (*б*) кремниевого транзистора

высоты потенциального барьера на коллекторном переходе. Поэтому при прямом смещении коллекторного перехода и соответственно при прямом смещении диода Шотки основная часть прямого тока коллектора будет проходить через диод Шотки. Этот ток связан с движением электронов из n-области коллектора в металлический электрод и не сопровождается инжекцией дырок в n-область коллектора. Таким образом, в высокоомной области коллектора практически не происходит накопления неосновных носителей заряда (рис. 7.9).

Кроме того, из-за меньшей высоты потенциального барьера на переходе Шотки по сравнению с высотой потенциального барьера на коллекторном переходе при тех же прямых токах коллектора на коллекторном переходе будет меньшее прямое напряжение, что соответствует меньшему количеству накопленных неосновных носителей заряда в базе транзистора при режиме насыщения (рис. 7.9). В результате время рассасывания в транзисторе с диодом Шотки оказывается значительно меньшим (несколько

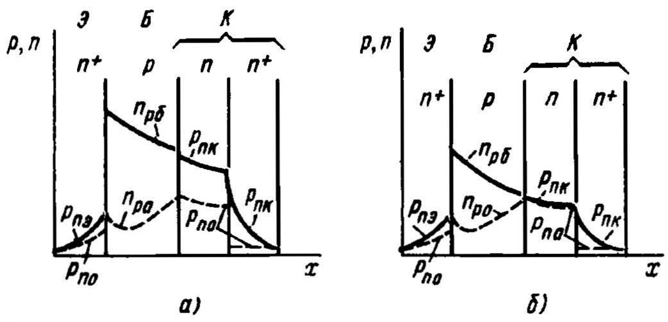


Рис. 7.9. Распределение неосновных носителей заряда в различных областях транзистора при его работе в режиме насыщения: *а* — в структуре обычного планарного транзистора; *б* — в структуре аналогичного транзистора с диодом Шотки, включенным параллельно коллекторному переходу

наносекунд), чем время рассасывания в транзисторе аналогичной структуры, но без шунтирующего диода Шотки.

Отметим, что на контакте алюминиевых электродов с сильнолегированными n^+-областями эмиттера и коллектора также могут возникать потенциальные барьеры, но их толщина оказывается настолько малой, что сквозь такие узкие потенциальные барьеры электроны могут проходить практически беспрепятственно путем туннелирования (см. рис. 7.8, *б*). Таким образом, на контактах алюминиевых электродов с эмиттерной областью и с сильнолегированной частью коллекторной области получаются омические переходы, а их формирование и формирование выпрямляющего перехода Шотки осуществляются во время одного процесса металлизации.

Изготовление интегрального транзистора с диодом Шотки не требует введения дополнительных технологических операций. Необходимо лишь изменить соответствующим образом фотошаблон, применяемый при проведении фотолитографии для снятия диоксида кремния под контакты, и расширить слой напыляемого алюминия за металлургическую границу коллекторного перехода. Однако при снятии диоксида кремния в месте выхода коллекторного перехода на поверхность монокристалла кремния и при обработке этой поверхности перед нанесением алюминиевой металлизации следует предотвратить возможность загрязнения *p-n*-перехода коллектора неконтролируемыми примесями.

Полевой транзистор с изолированным затвором

В связи с особенностями структуры МДП-транзисторов их можно формировать без специальных островков в монокристалле интегральной микросхемы, что упрощает технологию — уменьшает число технологических операций, удешевляет интегральные микросхемы и дает возможность увеличить плотность упаковки. Другая особенность и преимущество МДП-транзисторов в качестве активных элементов интегральных микросхем состоит в том, что при нулевом напряжении на затворе МДП-транзистора с индуцированным каналом ток стока практически отсутствует, т. е. мощность транзистором потребляется только во время подачи напряжения на затвор. Это уменьшение потребляемой мощности интегральных микросхем на МДП-транзисторах с индуцированным каналом особенно существенно для создания логических интегральных микросхем. Важным также является то обстоятельство, что цифровые интегральные микросхемы могут быть построены целиком на гальванически соединенных между собой МДП-транзисторах без использования других элементов.

Диод

Этот активный элемент широко используется в интегральных микросхемах, особенно в логических интегральных микросхемах. Для создания диода вообще достаточно сформировать только

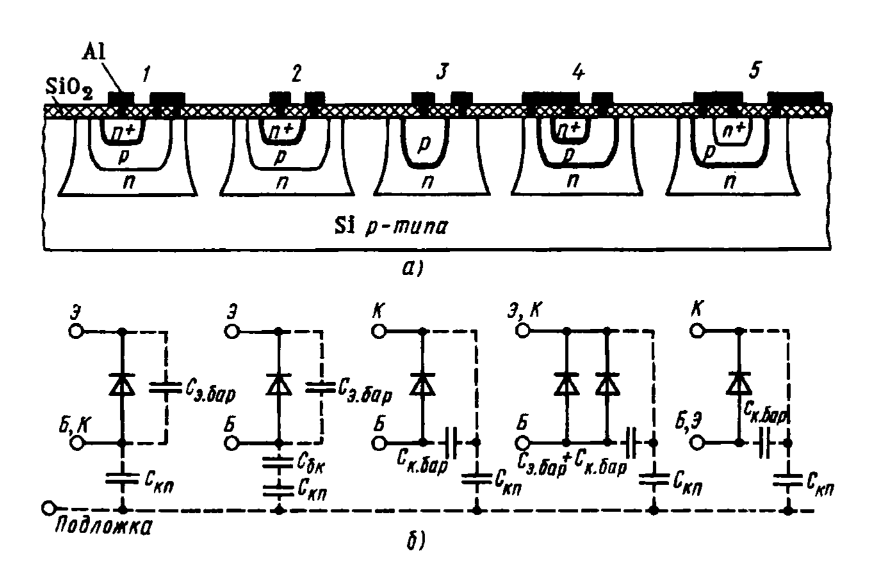


Рис. 7.10. Возможные варианты использования *p-n*-переходов транзисторной структуры в качестве диода *(а)* и их эквивалентные схемы *(б)*

один p-n-переход. Однако диодам в интегральных микросхемах придают транзисторную структуру и в зависимости от конкретного назначения используют тот или иной p-n-переход путем применения одного из пяти возможных вариантов включения (рис. 7.10).

В первом варианте (*1*) используется эмиттерный переход, а коллекторный короткозамкнут. Такое включение используют в цифровых микросхемах, так как в этом случае достигается наибольшее быстродействие: накопление носителей заряда может происходить только в базовой области, а она очень тонкая. Возможность накопления носителей заряда в коллекторной области исключена шунтированием коллекторного перехода. Время переключения может быть около 1 нс.

Во втором варианте (*2*) используется эмиттерный переход, а коллекторная цепь разомкнута.

В третьем варианте (*3*) используется коллекторный переход, а эмиттерной области при этом может и не быть, т. е. этап диффузии примесей для формирования эмиттерной области может быть исключен из технологического процесса. Если же эмиттерная область сформирована, то цепь эмиттера остается разомкнутой. Коллекторная область обычно является относительно высокоомной, поэтому такой диод имеет достаточно высокое пробивное напряжение (~50 В). Площадь коллекторного перехода значительно больше площади эмиттерного перехода, поэтому использование коллекторного перехода в качестве диодной структуры дает возможность пропускать бóльшие прямые токи.

В четвертом варианте (*4*) эмиттерную и коллекторную области соединяют между собой, т. е. эмиттерный и коллекторный переходы включают параллельно. Допустимый прямой ток при этом оказывается еще больше, но увеличивается также и суммарная барьерная емкость.

В пятом варианте (*5*) используется коллекторный переход, а эмиттерный короткозамкнут.

§ 7.5. ПАССИВНЫЕ ЭЛЕМЕНТЫ

Диффузионные резисторы

В полупроводниковых интегральных микросхемах биполярный транзистор является элементом с самой сложной структурой. Для его формирования необходимо провести последовательно несколько этапов диффузии примесей. Чтобы не усложнять технологию изготовления интегральной микросхемы, целесообразно для создания резисторов использовать одну из областей транзисторной структуры: эмиттер, базу или коллектор (рис. 7.11).

Эмиттерная область содержит наибольшую концентрацию примесей и обладает наименьшим удельным сопротивлением слоя. Поэтому эмиттерная область пригодна для формирования диффузионных резисторов только с малым сопротивлением (око-

ло 10 Ом). Из-за большой концентрации примесей температурные коэффициенты сопротивления таких резисторов будут малы (см. рис. 1.11).

Коллекторная область транзисторной структуры содержит наименьшую концентрацию примесей. Поэтому коллекторная область вообще пригодна для формирования диффузионных резисторов с большим сопротивлением, но из-за малой концент-

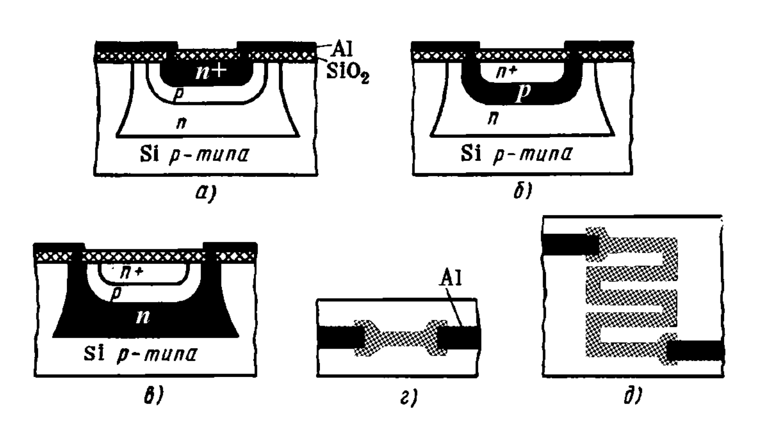


Рис. 7.11. Возможные варианты использования различных областей транзисторной структуры в качестве резистивных слоев для формирования диффузионных резисторов полупроводниковой интегральной микросхемы (*а, б, в*) и некоторые варианты конфигурации резистивного слоя и электродов резистора (*г, д*)

рации примесей температурные коэффициенты сопротивления таких резисторов велики.

Таким образом, для формирования диффузионных резисторов обычно используют базовую область транзисторной структуры. Без существенного увеличения площади, занимаемой диффузионным резистором, в базовой области могут быть созданы резисторы с сопротивлением до 50 кОм. В то же время такие диффузионные резисторы имеют приемлемые температурные зависимости сопротивления; во всяком случае, эти зависимости более слабые, чем для диффузионных резисторов, сформированных в коллекторной области.

Диффузионные резисторы, как и другие резистивные элементы, характеризуют следующие параметры: диапазон номинальных значений сопротивления, допуски по сопротивлению, температурный коэффициент сопротивления, допустимая мощность рассеяния и максимальное напряжение.

Для диффузионных резисторов *диапазон номинальных значений сопротивления,* как отмечалось, ограничен сверху.

В массовом производстве интегральных микросхем диффузионные резисторы не могут быть изготовлены с достаточно малыми пределами допусков по ряду причин. Например, трудно

выдерживать необходимую поверхностную концентрацию примесей и глубину диффузии с высокой точностью. Различные этапы фотолитографического процесса, несмотря на совершенство применяемого оборудования, вносят в абсолютное значение сопротивления диффузионного резистора существенную ошибку, особенно для резисторов с узкой шириной резистивной диффузионной полоски.

Температурный коэффициент сопротивления диффузионного резистора зависит от концентрации примесей в использованном диффузионном слое. Так как диффузионные резисторы формируют в базовых областях транзисторных структур, то концентрация примесей в этих областях определяется необходимыми параметрами и свойствами транзисторов.

Допустимая мощность рассеяния диффузионных резисторов ограничена малыми размерами и топологией резистивной диффузионной полоски резистора (рис. 7.11, *г, д*), а также связана с температурным коэффициентом сопротивления резистора, так как нагрев резистора проходящим током вызывает изменение сопротивления и приводит к нелинейности ВАХ.

Максимальное напряжение — это специфический параметр диффузионных резисторов. Диффузионный резистор, как отмечалось, представляет собой диффузионную полоску с определеным типом электропроводности, отделенную от подложки интегральной микросхемы *p-n*-переходом. Электронно-дырочный переход должен быть смещен в обратном направлении для изоляции диффузионной полоски от подложки. Поэтому максимальное падение напряжения на резисторе не может быть больше напряжения смещения, которое, в свою очередь, не может превышать пробивного напряжения изолирующего *p-n*-перехода.

Кроме перечисленных недостатков диффузионных резисторов и трудностей проектирования и создания интегральных микросхем с диффузионными резисторами необходимо отметить, что при формировании интегральных микросхем вообще и микросхем с диффузионными резисторами в частности в структуре интегральной микросхемы образуются паразитные элементы, которые могут нарушить нормальную работу интегральной микросхемы. Диффузионный резистор отделен от подложки *p-n*-переходом, который имеет собственную барьерную емкость. Эта емкость может оказывать существенное влияние на работу схемы при высоких частотах.

Другим паразитным элементом, возникающим возле диффузионного резистора, может быть паразитный биполярный транзистор, эмиттером которого является диффузионный резистивный слой, базой — коллекторная область исходной транзисторной структуры, коллектором — подложка интегральной микросхемы (рис. 7.11, *б*). Если изолирующий *p-n*-переход между резистивной полоской и подложкой (эмиттерный переход паразитного транзистора) окажется смещенным в прямом направлении, то паразитный транзистор может нарушить работу интегральной микро-

схемы. Поэтому-то и необходимо, чтобы падение иапряжения на диффузионном резисторе не превышало максимального напряжения.

Несмотря на отмеченные недостатки, диффузионные резисторы широко применяют в интегральиых микросхемах, так как их формироваиие ие требует дополнительных техиологических операций и ие удорожает схему.

Пленочные резисторы

Осиовой пленочного резистора является резистивная плеика из металла (хром, тантал, палладий), металлического сплава (нихром) или металлокерамики (кермет). Резистивиую пленку определенной конфигурации тем или ииым способом иаиосят на диэлектрическую подложку гибридной иитегральной микросхемы или иа окисленный кристалл полупроводииковой интегральной микросхемы, изготовленной по совмещениой технологии (рис. 7.12).

Материал, используемый для резистивных пленок, должен обеспечивать возможность получения широкого диапазона номииальных значений сопротивлеиия, обладать низким температурным коэффициентом сопротивлеиия, высокой коррозиоиной стойкостью и стабильностью параметров во времени. Наибольшее распространение получили нихромовые резистивные пленки из-за их стабильности и возможности их создания с малым температурным коэффициентом сопротивления.

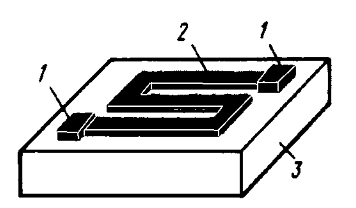


Рис. 7.12. Возможиая конфигурацня плеиочиого резнстора: 1 — металлическке электроды; 2 — резистивная пленка; 3 — диэлектрическая подложка гибридной ннтегральной микросхемы

После осаждения пленки нихрома на диэлектрическую подложку производят термообработку путем отжига пленки на воздухе. При этом нихромовая пленка покрывается слоем оксида, который значительно улучшает стабильность пленочного резистора.

При создаиии пленочных резисторов на диэлектрической подложке не образуется паразитных элементов, которые получаются в полупроводниковых интегральных микросхемах с диффузионными резисторами.

Для нанесения резистивиой пленки могут быть использованы различные способы: термическое испарение резистивного материала в вакууме (из резистивного испарителя или с помощью электронного луча), катодное распыление, анодирование осажденных на подложку металлических пленок и химическое осаждеиие при химической реакции в газовой фазе. Термическое испарение в вакууме — наиболее простой и производительный

метод получения резистивных пленок, а также других металлических и диэлектрических пленок, необходимых для изготовления интегральных микросхем. Унификация технологических операций получения различных элементов интегральной микросхемы, возможность изготовления этих различных элементов в едином технологическом цикле — это существенное преимущество термического испарения в вакууме среди других способов нанесения резистивных пленок.

Диффузионные конденсаторы

В качестве конденсаторов интегральных микросхем часто используют барьерную емкость *p-n*-перехода, смещенного в обратном направлении. Такой пассивный элемент интегральной микросхемы удобно формировать одновременно с формированием транзисторных структур или использовать непосредственно *p-n*-переходы транзисторных структур (рис. 7.13). Барьерная емкость *p-n*-перехода может быть использована как для создания конденсатора постоянной емкости, так и для конденсатора переменной емкости, которой можно управлять путем изменения постоянного смещения на переходе (см. § 3.31).

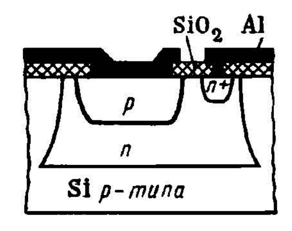


Рис. 7.13. Структура диффузионного конденсатора с использованием барьерной емкости коллекторного перехода транзисторной структуры (диффузия доноров для формирования эмиттерной области не проводилась)

Качество диффузионных конденсаторов, как и других конденсаторов, и их пригодность для применения в интегральных микросхемах необходимо оценивать по их технологической совместимости с технологией других (в первую очередь активных) элементов интегральных микросхем и по таким основным параметрам, как диапазон номинальных значений емкости (или удельная емкость), пробивное напряжение, добротность, допуски по емкости.

Диапазон номинальных значений емкости диффузионных конденсаторов, которые могут быть сформированы на отведенных для них площадях монокристалла полупроводника, определяется концентрацией примесей в прилегающих к переходу областях. Диффузионные конденсаторы, использующие эмиттерную емкость транзисторной структуры, имеют бóльшую удельную емкость по сравнению с конденсаторами на коллекторном переходе.

Однако при большой концентрации примесей в прилегающих к переходу областях и, следовательно, при малой толщине перехода будет малó пробивное напряжение такого перехода, а значит, и диффузионного конденсатора. Таким образом, удельную емкость и пробивное напряжение диффузионных конденсаторов надо рассматривать совместно. Взаимосвязь между этими пара-

метрами оказывается неблагоприятной для диффузионных конденсаторов.

Добротность — величина, обратная тангенсу угла диэлектрических потерь диффузионных конденсаторов, обычно значительно ниже добротности дискретных конденсаторов с диэлектрической изоляцией. Однако стоимость формирования диффузионных конденсаторов мала, так как они создаются в едином технологическом процессе с другими элементами интегральной микросхемы. Поэтому диффузионные конденсаторы широко используют в интегральных микросхемах, когда можно мириться с низкой добротностью.

Диффузионный конденсатор, как и другие элементы интегральной микросхемы, должен быть изолирован от остальных элементов и от подложки интегральной микросхемы. Часто эта изоляция осуществляется *p-n*-переходом. Поэтому при формировании диффузионного конденсатора одновременно образуется и структура паразитного транзистора, эмиттером которого является одна из областей (обкладок) диффузионного конденсатора, базой — другая область (обкладка), коллектором — подложка (рис. 7.13).

В связи с зависимостью барьерной емкости от напряжения смещения на *p-n*-переходе диффузионные конденсаторы могут быть использованы для усиления электромагнитных колебаний, т. е. могут быть активными элементами интегральных микросхем.

МДП-конденсаторы

В качестве диэлектрика такого конденсатора используют слой диоксида кремния, которым покрыт кристалл полупроводника (рис. 7.14). Одной обкладкой конденсатора является слой металла (обычно алюминия), нанесенный на поверхность слоя диоксида кремния одновременно с созданием межэлементных соединений и контактных площадок; другой обкладкой — сильнолегированная область полупроводника, которая формируется одновременно с формированием эмиттерных областей транзисторных структур интегральных микросхем. Таким образом, процесс изготовления МДП-конденсаторов также не требует проведения дополнительных операций для их формирования.

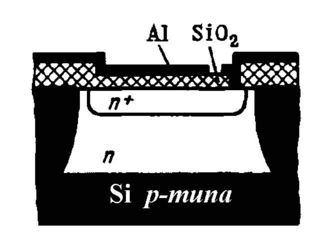


Рис. 7.14. Структура МДП-конденсатора

В островке, предназначенном для МДП-конденсатора, не формируют базовую область транзисторной структуры, т. е. не проводят диффузию примесей для создания базовой области. Поэтому под МДП-конденсатором есть только один *p-n*-переход между коллекторной областью транзис-

торной структуры и подложкой, который необходим для изоляции МДП-конденсатора от других элементов, расположенных на одной с ним полупроводниковой пластине.

Следует отметить, что емкость МДП-конденсатора может иметь сложную зависимость от напряжения постоянного смещения и от частоты переменного напряжения. Связано это с возможностью образования у поверхности полупроводника (в данном примере у поверхности n^+-области) обедненных и инверсных слоев под действием проникающего в полупроводник электрического поля. Частотная зависимость емкости МДП-конденсатора может появиться, если в полупроводнике образуется инверсный слой, в котором накопление и рассасывание неосновных для исходного полупроводника носителей заряда происходят в результате процессов тепловой генерации и рекомбинации, а инерционность этих процессов может быть велика.

Пленочные конденсаторы

Пленочные конденсаторы формируют на диэлектрической подложке гибридных интегральных микросхем. При этом необходимо провести по крайней мере три операции вакуумного напыления: нижней проводящей обкладки конденсатора, диэлектрической пленки и верхней проводящей обкладки (рис. 7.15). Такой пленочный конденсатор называют однослойным. Для получения большей емкости или для уменьшения площади, занимаемой конденсатором на подложке, можно делать многослойные пленочные конденсаторы, секции которых располагают «этажами» — одна над другой. Однако создание «многоэтажных» конденсаторов затрудняет процесс их изготовления, так как надо вводить дополнительные операции нанесения различных слоев, повышает стоимость, уменьшает надежность, увеличивает процент брака из-за увеличения краевого эффекта, уменьшения плотности и электрической прочности верхних диэлектрических слоев.

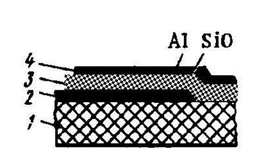


Рис. 7.15. Структура пленочного конденсатора:

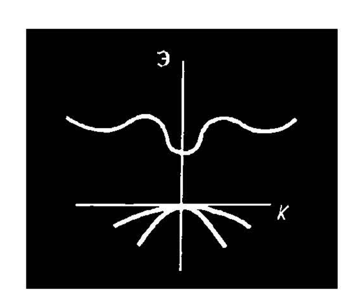
1 — диэлектрическая подложка; *2* — нижняя обкладка, *3* — диэлектрическая пленка; *4* — верхняя обкладка

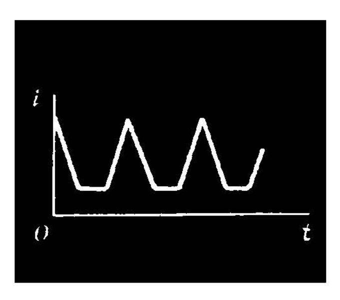
В качестве диэлектрика пленочных конденсаторов могут быть использованы различные материалы, но наиболее широко применяется монооксид кремния.

Контрольные вопросы

1. Что такое микроэлектроника?
2. Какие задачи стоят перед микроэлектроникой?
3. Что такое полупроводниковая интегральная микросхема?
4. Что такое гибридная интегральная микросхема?
5. Каково основное назначение цифровых и аналоговых интегральных микросхем?

6. Какими методами производят изоляцию элементов интегральных микросхем?
7. Какие элементы интегральных микросхем относят к активным?
8. Какие элементы интегральных микросхем считают пассивными?
9. Почему в качестве диодов интегральных микросхем используют транзисторные структуры?
10. Почему и как емкость диффузионного конденсатора зависит от напряжения постоянного смещения?
11. Какие недостатки и преимущества имеют диффузионные резисторы?
12. Почему не используют прямого смещения *p-n*-перехода диффузионного конденсатора?
13. Какие разновидности конденсаторов применяют в интегральных микросхемах?





Полупроводниковые приборы на эффекте междолинного перехода электронов

§ 8.1. ПРИНЦИП ДЕЙСТВИЯ ГЕНЕРАТОРОВ ГАННА

Генератор Ганна — это полупроводниковый прибор, действие которого основано на появлении отрицательного дифференциального сопротивления под воздействием сильного электрического поля, предназначенный для генерации и усиления СВЧ-колебаний.

Впервые генерацию электромагнитных СВЧ-колебаний при приложении постоянного напряжения к кристаллу однородного арсенида галлия или фосфида индия наблюдал в 1963 г. американский ученый Дж. Б. Ганн. Поэтому в советской технической литературе подобные приборы называют генераторами Ганна, или диодами Ганна, хотя в их структуре нет выпрямляющего электрического перехода. В зарубежной литературе чаще используют сокращенное наименование TED (*Transferred — Electron Devices*).

Физические основы отрицательного дифференциального сопротивления при междолинном переходе носителей заряда

Энергетическая диаграмма некоторых полупроводников (например, арсенида галлия), построенная в пространстве квазиимпульсов (в k-пространстве), может иметь несколько минимумов (см. рис. 1.15). В таком полупроводнике могут существовать электроны с разными подвижностями — «легкие» и «тяжелые» (см. § 1.10). Соотношение между концентрациями «легких» n_1 и «тяжелых» n_2 электронов изменяется при изменении напряжен-

ности электрического поля, так как в сильном электрическом поле (при напряженности, большей порогового значения $E > E_{пор}$) электроны, приобретая дополнительную энергию, превышающую $\Delta Э_1$ (см. рис. 1.15), переходят в боковые долины и становятся «тяжелыми». Если при этом еще не происходит заметной ударной ионизации, то общая концентрация электронов остается неизменной и равной равновесной концентрации: $n_1 + n_2 = n_0$.

Обозначив подвижность «легких» электронов μ_1, подвижность «тяжелых» электронов μ_2, выражение для плотности тока через кристалл полупроводника запишем так:

$$J = q(n_1\mu_1 + n_2\mu_2)E .$$

При слабых электрических полях ($E < E_{пор}$) практически все электроны находятся в центральной долине, $n_1 \approx n_0$ и плотность тока при этом $J = qn_0\mu_1 E$, что соответствует участку *1* ВАХ кристалла полупроводника (рис. 8.1). При сильных электрических полях ($E \gg E_{пор}$) можно предположить, что практически все электроны приобретут добавочную энергию, бо́льшую $\Delta Э_1$, и окажутся в боковой долине. В этом случае $n_2 \approx n_0$ и $J = qn_0\mu_2 E$, что соответствует участку *3* ВАХ (рис. 8.1). При средних напряженностях электрического поля, лишь немного превышающих пороговую напряженность, плотность тока определяется соотношением концентрации «легких» и «тяжелых» электронов (участок *2* ВАХ на рис. 8.1).

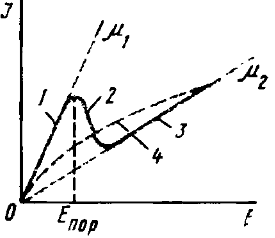


Рис. 8.1. Зависимость плотности тока через полупроводник с многодолинной структурой зоны проводимости от напряженности электрического поля:

1 — при наличии электронов только с большой подвижностью μ_1 (в центральной долине); *2* — переходный участок; *3* — при наличии электронов только с малой подвижностью $\mu_2 \ll \mu_1$ (в боковой долине); *4* — при наличии в кристалле локальных напряженностей, отличных от средней напряженности электрического поля

Для возникновения отрицательного дифференциального сопротивления необходим одновременный переход большинства электронов из центральной долины в боковую при пороговой напряженности электрического поля. Но практически получить статическую ВАХ, соответствующую сплошной кривой на рис. 8.1, не удается, так как в кристалле или около невыпрямляющих контактов всегда есть неоднородности, в результате чего возникают локальные напряженности электрического поля, превышающие среднюю напряженность. Превращение в этих местах «легких» электронов в «тяжелые» еще больше увеличивает неоднородность электрического поля. Поэтому практически не получается одновременного перехода большинства электронов в кристалле из центральной долины в боковую и статическая ВАХ получается без участка с отрицательным дифференциальным сопротивлением (штриховая кривая *4* на рис. 8.1).

Эффект Ганна

Пусть на однородно легированный кристалл арсенида галлия (рис. 8.2), имеющий два невыпрямляющих электрических перехода с электродами катода и анода, подано постоянное напряжение, создающее в кристалле напряженность электрического поля, несколько меньшую пороговой напряженности. При этом все свободные электроны в кристалле являются «легкими» и плотность тока через кристалл имеет максимальное значение:

$$J_{max} = qn_0\mu_1 E_0 = qn_0 v_0 .$$

Локальная напряженность электрического поля около невыпрямляющих контактов из-за наличия различных дефектов может превышать пороговую напряженность электрического поля.

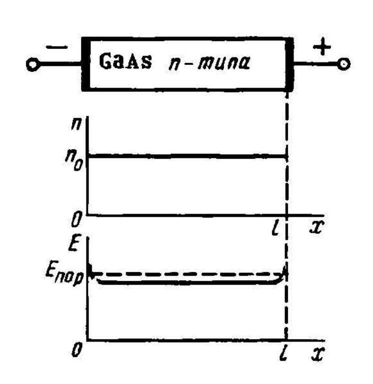


Рис. 8.2. Распределение концентрации электронов и напряженности электрического поля в кристалле генератора Ганна в первый момент после подачи постоянного напряжения

Рис. 8.3. Распределение концентрации электронов и напряженности электрического поля в кристалле после формирования домена

Это обеспечит образование «тяжелых» электронов около катода, которые, двигаясь относительно медленно к аноду, создают отрицательный заряд. «Легкие» электроны в остальной части кристалла движутся к аноду быстрее «тяжелых». Поэтому около пакета «тяжелых» электронов со стороны анода получается недостаток электронов, что равносильно образованию некоторого положительного заряда, состоящего из нескомпенсированных ионизированных доноров (рис. 8.3). Таким образом образуется *домен*, состоящий из двух слоев: слой со стороны катода из-за избытка «тяжелых» электронов имеет отрицательный заряд, слой со стороны анода из-за недостатка электронов имеет положительный заряд. Домен обладает своим электрическим полем $E_{дом}$, направленным в ту же сторону, что и поле, созданное внешним

напряжением. В результате по мере образования домена поле в нем растет, а за пределами домена уменьшается, т. е. скорость движения «тяжелых» электронов внутри домена увеличивается, а скорость движения «легких» электронов за пределами домена уменьшается. В некоторый момент времени скорость движения «тяжелых» электронов (скорость домена) оказывается равной скорости движения «легких» электронов: $v_1 = v_2$, или $\mu_1 E_1 = \mu_2 E_2$, где v_1 — скорость движения электронов за пределами домена; v_2 — скорость движения электронов внутри домена, что соответствует скорости движения домена от катода к аноду.

Очевидно, что $v_1 < v_0$, так как $E_1 < E_0$. Поэтому после образования домена плотность тока через кристалл уменьшится до

$$J_{\min} = q n_0 v_1 .$$

Минимальное значение плотности тока через кристалл будет сохраняться в течение всего времени движения домена через кристалл или в течение времени пролета

$$t_{\text{прол}} = l / v_2 ,$$

где l — длина кристалла.

При достижении анода домен исчезает и плотность тока возрастает до значения $J_{\max}$, соответствующего отсутствию домена. Сразу после этого у катода формируется новый домен, и процесс повторяется. На рис. 8.4 показана зависимость тока, проходящего через кристалл, от времени.

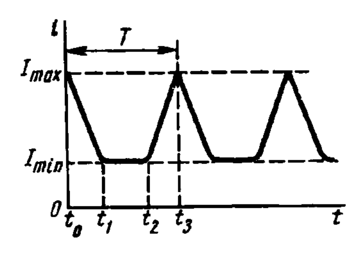


Рис. 8.4. Зависимость тока, проходящего через генератор Ганна, от времени:

t_0 — момент времени, соответствующий началу образования домена; t_1 — момент времени, соответствующий окончанию образования домена; t_2 — момент времени, соответствующий началу исчезновения домена на аноде; t_3 — момент времени, соответствующий полному исчезновению домена на аноде и зарождению второго домена на катоде

Рассмотренный механизм действия прибора с междолинным переходом электронов соответствует пролетному режиму работы. В этом режиме работы, как было отмечено, электрическое поле в домене растет во время его формирования при одновременном уменьшении напряженности электрического поля за пределами домена. По этой причине в кристалле может образоваться только один домен, так как переход электронов из центральной долины в боковую может происходить только в домене, где суммарная напряженность электрического поля превышает пороговое значение.

Время формирования домена определяется временем максвелловской релаксации ($\tau = \varepsilon\varepsilon_0\varrho$). Время пролета домена от катода до анода должно быть больше времени его формирования. Поэтому условие возникновения колебаний тока в генераторе Ганна можно сформулировать следующим образом:

$$t_{\text{прол}} = l/v > \varepsilon\varepsilon_0\varrho, \text{ или } n_0 l \gg \varepsilon\varepsilon_0 v/(q\mu_2) .$$

При скорости иасыщеиия для электронов в арсениде галлия $v \approx 10^7$ см/с правая часть неравенства имеет значение около 10^{11} см$^{-2}$.

§ 8.2. ТЕХНОЛОГИЯ ИЗГОТОВЛЕНИЯ ГЕНЕРАТОРОВ ГАННА

Несмотря иа приицнпиальную простоту коиструкции генератора Ганиа, представляющего собой однородный кристалл арсенида галлия с электропроводиостью n-типа с двумя контактами, нанесенными иа противоположиые стороны кристалла, в производстве таких приборов встречаются трудностн, связанные прежде всего с необходимостью создания тонких слоев однородно легированного арсеиида галлия (или другого полупроводникового материала).

В генераторах Ганна, имеющих толщину кристалла с высоким удельным сопротивлением более 200 мкм и работающих в пролетном режиме, колебаиия получаются, как правило, некогерентными. Это вызвано наличием в каждом из кристаллов нескольких дефектов, на которых могут зарождаться домены. Путь, пробегаемый доменом от места его зарождения до анода, определяет период колебаний. Поэтому если домены зарождаются на различных иеодиородиостях кристалла, т. е. на разлнчных расстояниях от аиода, то колебаиия будут иметь шумовой характер. Примеиеиие приборов с междолиным переходом электронов практически оправдаио в диапазоие частот более 1 ГГц, что соответствует толщиие высокоомного полупроводника $l \leqslant 100$ мкм.

В свою очередь, наименьшая толщина кристаллов полупроводника с высоким удельным сопротивлением для генераторов Гаииа кроме технологических трудностей ограничена тем, что для перехода электрона из цеитральной в боковую долину необходима дополнительная энергия, которую электрон может приобрести в электрическом поле после прохождения некоторого расстояния. Так, дополиительную энергию (0,3 эВ) электрон может приобрести в арсениде галлия при пороговой напряженности электрического поля, равной 3 кВ/см, только на расстоянии 1 мкм. Сам домен также имеет иекоторую протяжениость, определяемую в основном толщиной слоя положительной части домена, так как плотность положительного заряда в обеднением электронами слое ограничена концентрацией доноров в высокоомном полупроводнике ($N_д = 10^{14} ... 10^{16}$ см$^{-3}$). Ограничение минимальной толщины высокоомного кристалла приводит к ограничеиию максимальной частоты генерации в пролетном режиме для генераторов Ганна из арсеиида галлия значениями около 50 ГГц.

Первые приборы на эффекте междолинного перехода электронов изготовляли из арсенида галлия н фосфида индия со сплавными омическими переходами между кристаллом полупроводника и электродами. Но методом вплавления трудно получнть с вос-

производимыми результатами малую толщину полупроводника между омическими переходами. Поэтому в настоящее время для этих целей используют эпитаксиальные слои, наносимые на сильнолегированную подложку с электропроводностью n-типа. Многослойные структуры n^+-n-n^+ удобны, во-первых, при создании невыпрямляющих электрических переходов между металлическими электродами и сильнолегированными слоями полупроводника. Во-вторых, пластина с такой структурой достаточно механически прочна при резке ее на отдельные кристаллы, при пайке выводов и монтаже в корпус.

§ 8.3. ПАРАМЕТРЫ И СВОЙСТВА ГЕНЕРАТОРОВ ГАННА

Как любой генератор СВЧ-диапазона, генератор Ганна характеризуется генерируемой мощностью (при импульсной и непрерывной работе), длиной волны или частотой генерируемых колебаний, коэффициентом полезного действия, уровнем частотных и амплитудных шумов и другими параметрами.

Выходная непрерывная мощность генераторов Ганна в пролетном режиме обычно составляет десятки — сотни милливатт, а при импульсной работе достигает сотен ватт.

Рабочая частота в пролетном режиме обратно пропорциональна длине или толщине высокоомной части кристалла ($f = v/l$). Связь между генерируемой мощностью и частотой можно представить в виде

$$P = U^2/z = E^2 l^2/z = \frac{E^2 v^2}{z f^2} \sim \frac{1}{f^2}.$$

Отсюда следует, во-первых, что мощность генерируемых СВЧ-колебаний зависит от полного сопротивления z или от площади рабочей части высокоомного слоя полупроводника. Во-вторых, приведенное соотношение указывает на то, что ожидаемое изменение мощности с частотой пропорционально $1/f^2$. Усредненные зависимости генерируемых мощностей от частоты различных генераторов Ганна из арсенида галлия и из фосфида индия приведены на рис. 8.5.

Генераторы Ганна из арсенида галлия могут генерировать СВЧ-колебания от 1 до 50 ГГц (каждый прибор рассчитан на свою частоту). Несколько бóльшие частоты получены на генераторах Ганна из фосфида индия в связи с бóльшими значениями максимальных скоростей электронов, но качество приборов из этого материала значительно ниже из-за недостаточной отработки технологии изготовления материала. Преимущество фосфида индия перед арсенидом галлия как исходного материала для приборов на эффекте междолинного перехода электронов — большее значение пороговой напряженности электрического поля (10,5 и 3,2 кВ/см соответственно). Это отличие должно привести к созданию из фосфида индия генераторов Ганна со значительно

большими выходными мощностями. Для создания генераторов Ганна с еще большими частотами генерируемых колебаний, но с меньшими мощностями представляют интерес тройные соединения Ga In Sb, так как в них дрейфовые скорости электронов велики, но меньше пороговые напряженности электрического поля.

Коэффициент полезного действия генераторов Ганна может быть различным (от 1 до 30%), так как существенно отличаются технологии изготовления приборов и качество исходного полупроводникового материала.

В связи с возможным наличием в кристалле генератора Ганна нескольких неоднородностей зарождение домена, как отмечалось в § 8.2, может происходить в различные моменты времени на разном расстоянии от анода. Поэтому частота колебаний будет изменяться, т. е. могут возникать *частотные шумы*. Кроме частотных шумов в генераторах Ганна существуют *амплитудные шумы*, основной причиной которых являются флуктуации в скоростях движения электронов. Обычно амплитудные шумы в генераторах Ганна малы, так как дрейфовая скорость в сильных электрических полях, существующих в этих приборах, насыщена и слабо изменяется при изменении электрического поля. Шумы генераторов Ганна значительно меньше шумов лавинно-пролетных диодов.

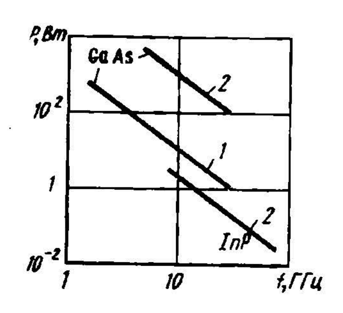


Рис. 8.5. Усредненные для многих генераторов Ганна зависимости генерируемой СВЧ-мощности от частоты при непрерывной работе *(1)* и при работе в импульсном режиме *(2)*

Важным для практического применения генераторов Ганна является вопрос о возможности их *частотной перестройки* в достаточно широком диапазоне. Из принципа действия генератора Ганна ясно, что частота его должна слабо зависеть от приложенного напряжения. С увеличением приложенного напряжения несколько возрастает толщина домена, а скорость его движения изменяется незначительно. В результате при изменении напряжения от порогового до пробивного частота колебаний увеличивается всего на десятые доли процента.

Срок службы генераторов Ганна относительно мал, что связано с одновременным воздействием на кристалл полупроводника таких факторов, как сильное электрическое поле и перегрев кристалла из-за выделяющейся в нем мощности.

§ 8.4. ГЕНЕРАТОРЫ С ОГРАНИЧЕНИЕМ НАКОПЛЕНИЯ ОБЪЕМНОГО ЗАРЯДА

Генератор с ограничением накопления объемного заряда — это генератор Ганна, работающий в режиме неполного формирования домена.

Предположим, что кроме постоянного смещения, большего $E_{пор}l$, где $E_{пор}$ — пороговая напряженность поля, к прибору приложено также синусоидальное напряжение достаточно большой амплитуды (рис. 8.6). В этом случае часть периода напряжение на приборе будет меньше порогового значения $E_{пор}l$.

Пока мгновенные значения напряжения в приборе больше, чем пороговое напряжение $E_{пор}l$, у катода формируется домен. Однако если прибор поместить в резонатор, настроенный на достаточно большую частоту, то домен, не успев сформироваться, начнет рассасываться, так как суммарное напряжение смещения и колебаний резонатора в следующую часть периода будет меньше, чем пороговое напряжение $E_{пор}l$. Во время формирования домена ток, проходящий через прибор, уменьшается, а во время рассасывания возрастает. Таким образом, возникают периодические колебания тока, но период этих колебаний определяется уже не пролетным временем домена, а частотой резонатора. Такие генераторы называют *генераторами с ограничением накопления объемного заряда* (ОНОЗ). Этот режим работы генератора Ганна [в английской транскрипции — *LSA* (*Limited Spacecharge Accumulation*) был обнаружен в 1966 г.

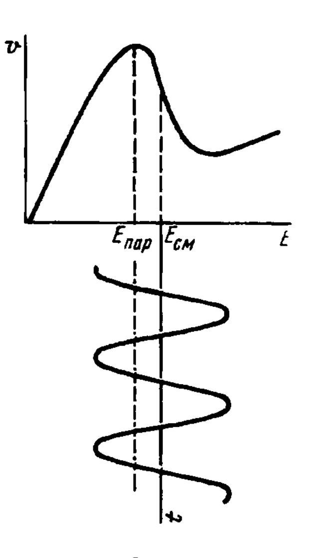


Рис. 8.6. Пояснение принципа действия генератора с ограничением накопления объемного заряда

Для установления режима ОНОЗ необходимо выполнение ряда условий. Прежде всего нужно, чтобы домен не успел сформироваться за то время, пока напряжение на генераторе превышает пороговое значение $E_{пор}l$, период колебаний T должен быть меньше $(2...3)\tau'$, где $\tau' = \varrho'\varepsilon\varepsilon_0$ — максвелловское время релаксации для материала в состоянии с отрицательным дифференциальным сопротивлением.

За то время, пока мгновенные значения напряжения на генераторе меньше порогового напряжения $E_{пор}l$, возникший у катода домен должен успеть полностью исчезнуть. Для выполнения этого условия необходимо, чтобы период колебаний был много больше максвелловского времени релаксации τ для материала, находящегося в слабом электрическом поле. Таким образом, для существования режима ОНОЗ необходимо выполнение условий

$$T=\frac{1}{f}<(2...3)\tau'=\frac{(2...3)\varepsilon\varepsilon_0}{qn_0\mu'} \text{ и } T\gg\tau=\varrho\varepsilon\varepsilon_0=\frac{\varepsilon\varepsilon_0}{qn_0\mu_1}$$

или

$$\frac{\varepsilon\varepsilon_0}{q\mu_1}\ll\frac{n_0}{f}<\frac{(2...3)\varepsilon\varepsilon_0}{q\mu'},$$

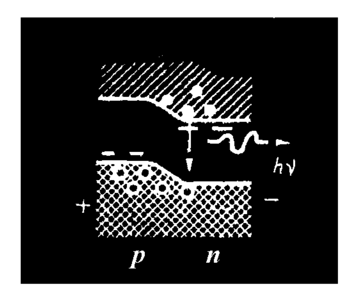
где μ_1 и μ' — подвижности электронов в слабом поле и при наличии отрицательного дифференциального сопротивления.

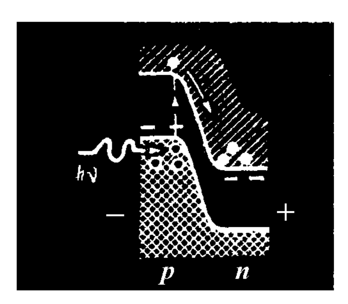
Одним из преимуществ генераторов с ОНОЗ перед генераторами Ганна является независимость частоты генерируемых колебаний от толщины кристалла. Другое, более важное, их преимущество — возможность получения бо́льших мощностей генерации на более высоких частотах. Генераторы с ОНОЗ могут быть изготовлены из относительно толстых пластинок высокоомного арсенида галлия, что позволяет подавать на такие приборы значительные напряжения.

С помощью генераторов с ОНОЗ можно достичь мощности непрерывных колебаний порядка десятых долей ватта в диапазоне 30...300 ГГц с КПД, равным 20%, и сотен ватт в импульсном режиме.

Контрольные вопросы

1. Что такое генератор Ганна?
2. Почему на статической ВАХ не получается участка с отрицательным дифференциальным сопротивлением, если энергетическая диаграмма полупроводника имеет несколько минимумов и в полупроводнике могут существовать «легкие» и «тяжелые» электроны?
3. Что представляет собой домен, возникающий в кристалле полупроводника при переходе электронов из центральной в боковую долину энергетической диаграммы?
4. Почему в кристалле полупроводника может возникнуть только один домен?
5. Чем определяется частота электрических колебаний, возникающих в генераторе Ганна?
6. Чем отличается принцип действия генераторов с ограничением накопления объемного заряда от принципа действия генераторов Ганна?
7. Какими свойствами должны обладать кристаллы полупроводника, предназначенные для изготовления генераторов Ганна?





Оптоэлектронные полупроводниковые приборы

§ 9.1. КЛАССИФИКАЦИЯ ОПТОЭЛЕКТРОННЫХ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

Оптоэлектронный полупроводниковый прибор — это полупроводниковый прибор, излучающий или преобразующий электромагнитное излучение, чувствительный к этому излучению в видимой, инфрактрасной и (или) ультрафиолетовой областях спектра или использующий подобное излучение для внутреннего взаимодействия его элементов.

Оптическим диапазоном электромагнитных колебаний, для которого предназначены оптоэлектронные полупроводниковые приборы, условно считают диапазон длин волн от 1 мм до 1 нм (рис. 9.1).

Оптоэлектронные полупроводниковые приборы можно подразделить на полупроводниковые излучатели, приемники излучения, оптопары и оптоэлектронные интегральные микросхемы (рис. 9.2).

Полупроводниковый излучатель — это оптоэлектронный полупроводниковый прибор, преобразующий электрическую энергию в энергию электромагнитного излучения в видимой, инфракрасной и ультрафиолетовой областях спектра.

Многие полупроводниковые излучатели могут излучать только некогерентные электромагнитные колебания. К ним относятся полупроводниковые излучатели видимой области спектра — полупроводниковые приборы отображения информации (светоизлучающие диоды, полупроводниковые знаковые индикаторы, шкалы и экраны), а также полупроводниковые излучатели инфракрасной области спектра — инфракрасные излучающие диоды.

Когерентные полупроводниковые излучатели — это полупроводниковые лазеры с различными видами возбуждения. Они могут излучать электромагнитные волны с определенной амплитудой, частотой, фазой, направлением распространения и поляризацией, что и соответствует понятию когерентности.

Полупроводниковый приемник излучения — это оптоэлектронный полупроводниковый прибор, чувствительный к электромагнитному излучению видимой, инфракрасной и (или) ультрафиолетовой области спектра или преобразующий энергию электромагнитного излучения непосредственно в электрическую энергию.

К полупроводниковым приемникам излучения относятся фоторезисторы, фотодиоды, фотоэлементы, фототранзисторы и фототиристоры.

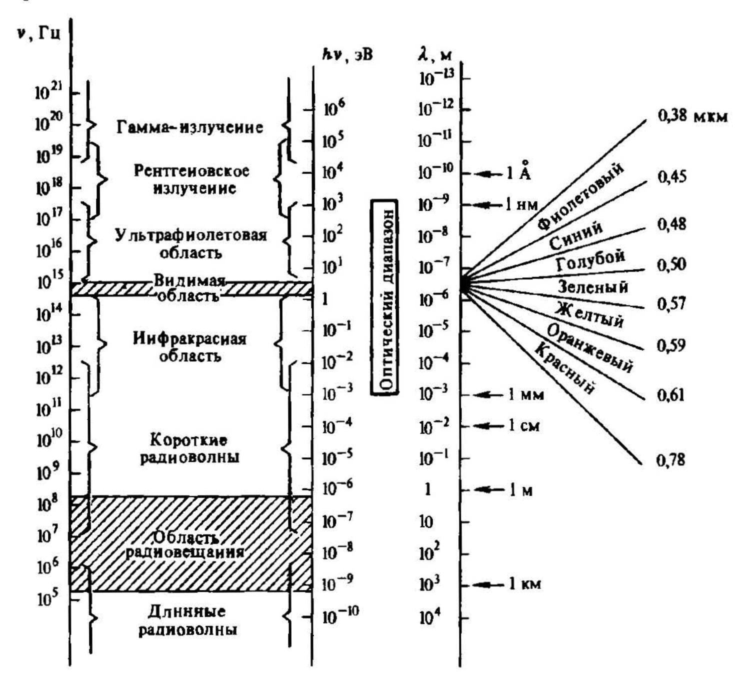


Рис. 9.1. Шкала электромагнитных волн

§ 9.2. ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ ОТОБРАЖЕНИЯ ИНФОРМАЦИИ И ИНФРАКРАСНЫЕ ИЗЛУЧАЮЩИЕ ДИОДЫ

Полупроводниковый прибор отображения информации — это полупроводниковый излучатель энергии видимой области спектра, предназначенный для отображения визуальной информации.

Инфракрасный излучающий диод (ИК-диод) — это полупроводниковый излучатель, представляющий собой диод, способный излучать энергию инфракрасной области спектра.

Принцип действия

Рассмотрим полупроводниковые приборы отображения информации, структура которых имеет выпрямляющий электрический переход (*p-n*-переход или гетеропереход), а также инфракрасные

излучающие диоды, так как принцип действия этих приборов один и тот же. Электролюминесцентные порошковые и пленочные излучатели хотя и относятся к полупроводниковым приборам отображения информации (см. рис. 9.2), но отличаются по структуре и принципу действия. Поэтому они будут рассмотрены в последующих параграфах.

Излучение полупроводниковых приборов отображения информации и инфракрасных излучающих диодов вызвано самопроизвольной рекомбинацией носителей заряда при прохождении прямого тока через выпрямляющий электрический переход. Рекомбинация носителей заряда может происходить как в самом выпрямляющем электрическом переходе, так и в прилегающих к этому переходу областях структуры. Чтобы кванты энергии — фотоны, освободившиеся при рекомбинации, соответствовали

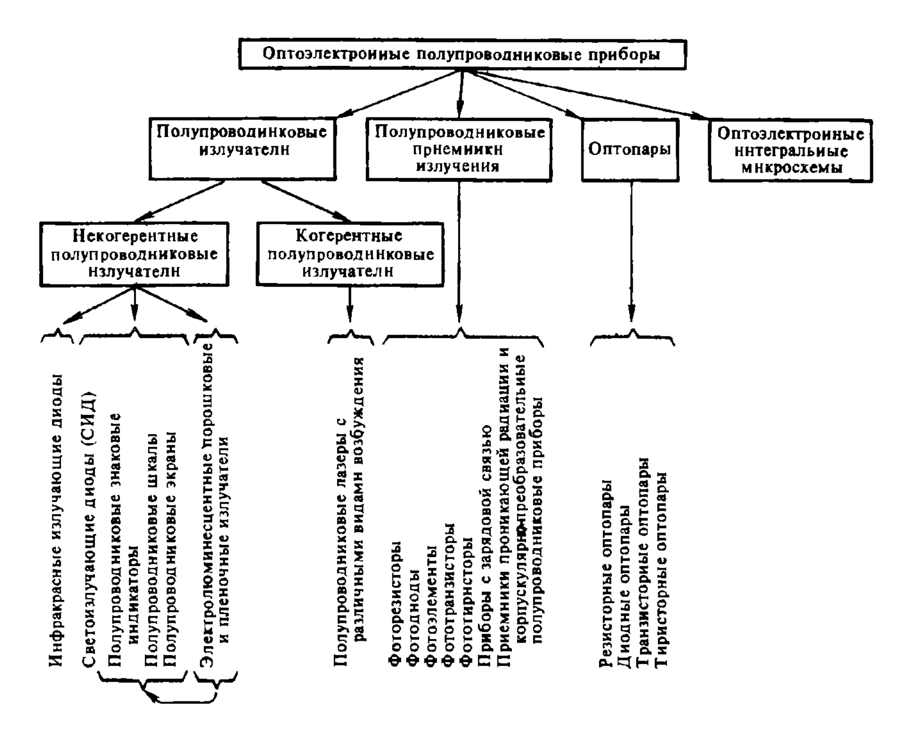


Рис. 9.2. Классификация оптоэлектронных полупроводниковых приборов

квантам видимого света, ширина запрещенной зоны исходного полупроводника должна быть относительно большой ($\Delta Э > 1,7$ эВ). При меньшей ширине запрещенной зоны исходного полупроводника кванты энергии, освобождающиеся при рекомбинации носителей заряда, соответствуют инфракрасной области излучения. Таким образом, разница между полупроводниковыми приборами отображения информации (визуального представления информа-

ции) и инфракрасными излучающими диодами заключается только в отличии исходного полупроводникового материала.

Если бы рекомбинация неравновесных электронов и дырок, вводимых в выпрямляющий электрический переход или в прилегающие к нему области при прохождении прямого тока, происходила только с излучением фотонов, то внутренний квантовый выход (отношение излученных фотонов к числу рекомбинировавших пар носителей) был бы равен 100%. Однако значительная часть актов рекомбинации может заканчиваться выделением энергии в виде элементарных квантов тепловых колебаний — фононов. Такие переходы электронов между энергетическими уровнями называют безызлучательными. Соотношение между излучательными и безызлучательными переходами зависит от ряда причин, в частности от структуры энергетических зон полупроводника, наличия примесей, которые могут увеличивать или уменьшать вероятность излучательных переходов. Из освоенных в настоящее время полупроводниковых материалов наилучшими с точки зрения внутреннего квантового выхода являются соединения $GaAs_{1-x}P_x$ при $x = 0...0,45$. Ширина запрещенной зоны этих соединений увеличивается от 1,424 при $x = 0$ до 1,977 эВ при $x = 0,45$.

В полупроводниковых излучателях из арсенида галлия, т. е. при $x = 0$ в указанной системе соединений, внутренний квантовый выход достигает значений, близких к 100%. При использовании других полупроводниковых материалов внутренний квантовый выход составляет иногда только единицы процентов, но и при таких значениях излучение оказывается достаточным для практического использования.

Конструкция и технология изготовления

Даже при высоком внутреннем квантовом выходе внешний квантовый выход полупроводниковых излучателей оказывается значительно меньшим из-за поглощения фотонов в полупроводнике до выхода их в окружающее пространство и из-за потерь при полном внутреннем отражении фотонов, падающих на границу

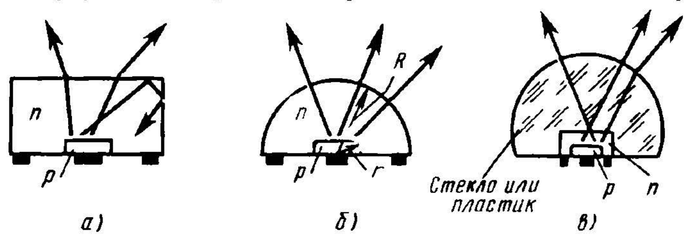


Рис. 9.3. Структуры некогерентных полупроводниковых излучателей:

а — плоская планарная; *б* — полусферическая; *в* — плоская планарная с прозрачным полусферическим покрытием

раздела полупроводника и окружающей атмосферы под углом, превышающим критический угол полного внутреннего отражения $\varphi_{кр} = \arcsin(1/n_r)$, где n_r — коэффициент преломления полупроводника. Обычно $\varphi_{кр} \leqslant 17°$.

Таким образом, из полупроводникового излучателя, имеющего простейшую плоскую структуру (рис. 9.3, *а*), в окружающее пространство выходит только часть фотонов, возникших в выпрямляющем электрическом переходе или вблизи него. Внешний квантовый выход удается увеличить при использовании более сложных конструкций полупроводниковых излучателей со структурой в виде полусферического кристалла полупроводника (рис. 9.3, *б*) или плоской структуры с прозрачным полусферическим покрытием (рис. 9.3, *в*). В полупроводниковых излучателях с полусферической структурой для всей поверхности угол падения фотонов оказывается меньше критического угла полного внутреннего отражения при большом отношении радиусов R и r (рис. 9.3, *б*), т. е. при

$$R/r > n_r/n_{rср},$$

где n_r и $n_{rср}$ — коэффициенты преломления электромагнитного излучения в полупроводнике и среде, окружающей полупроводниковый излучатель ($n_{rср} = 1$ для воздуха).

Однако в полупроводниковых излучателях с полусферической структурой несколько возрастают потери фотонов в результате поглощения, так как увеличивается длина их пути от места возникновения до поверхности кристалла. Все полупроводниковые излучатели с полусферической структурой имеют внешний квантовый выход на порядок выше, чем у излучателей с плоской конструкцией.

Значительно проще технология изготовления полупроводниковых излучателей с прозрачным полусферическим (или параболическим) покрытием из различных пластических материалов с высоким коэффициентом преломления для увеличения критического угла полного внутреннего отражения в полупроводнике.

Основным методом формирования *p-n*-переходов и гетеропереходов при создании полупроводниковых излучателей на основе арсенида галлия GaAs, фосфида галлия GaP, твердых растворов этих соединений $GaAs_{1-x}P_x$ и других соединений типа $A^{III}B^{V}$ является метод эпитаксиального наращивания. Обычно это жидкофазная эпитаксия, иногда — эпитаксия из газовой фазы.

Для формирования *p-n*-переходов в карбиде кремния SiC используют метод диффузии примесей, а иногда и метод эпитаксиального наращивания. Одной из интересных особенностей карбида кремния является его политипизм, т. е. существование нескольких кристаллических модификаций, отличающихся, в частности, шириной запрещенной зоны. Высокая физико-химическая стабильность карбида кремния и относительно малые коэффициенты диффузии примесей в нем создают предпосылки для изготовления высокостабильных полупроводниковых излучателей

на основе этого материала. Однако технология изготовления монокристаллов карбида кремиия и технология формироваиия выпрямляющих электрических переходов в этих монокристаллах отличаются сложностью. Кроме того, в полупроводниковых излучателях из карбида кремния не удается получить высокий квантовый выход.

Большой интерес для изготовления полупроводниковых излучателей представляет нитрид галлия GaN, имеющий наибольшую шириу запрещенной зоиы ($\Delta Э = 3{,}5$ эВ) среди соединений типа $A^{III}B^{V}$, освоениых в технологическом отношении. Энергии фотоиов, которые могут быть возбуждены в этом материале, перекрывают всю видимую область спектра. Однако независимо от метода получения и легирования нитрид галлия обладает только электропроводиостью n-типа. Поэтому для получения излучения при рекомбинации неравновесных иосителей заряда в данном случае надо создать выпрямляющий электрический переход в виде перехода Шотки на контакте металла с нитридом галлия.

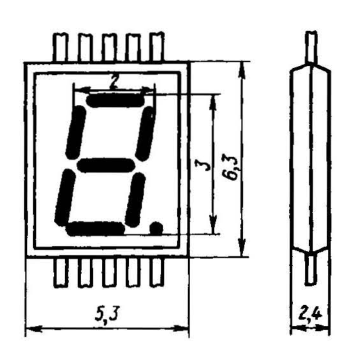


Рис. 9.4. Конструкция одноразрядного знакового индикатора (цнфрового индикатора)

Сопоставление эффективности люмиисценции различных материалов показывает, что кваитовый выход растет с увеличением длииы волиы. Поэтому, если зрительное восприятие информации не является обязательным, предпочтение следует отдать инфракрасным излучательным диодам на осиове арсенида галлия.

Полупроводниковые приборы отображения информации в зависимости от структуры, конструкции и, конечно, назиачения могут быть разделены иа светоизлучающие диоды, полупроводииковые знаковые индикаторы, шкалы и экраны (см. рис. 9.2).

Светоизлучающий диод (СИД) — это полупроводниковый прибор отображения информации, представляющий собой диод.

Таким образом, в структуре светоизлучающего диода есть только один выпрямляющий электрический переход (рис. 9.3) или один полупроводниковый излучающий элемент.

Полупроводниковый излучающий элемент — часть полупроводникового прибора отображения информации, состоящая из излучающей поверхиости и контактов для подключения к электрической схеме.

Полупроводниковый знаковый индикатор — это полупроводниковый прибор отображения информации, состоящий из полупроводниковых излучающих элементов, предиазначенный для представления информации в виде знаков и организованный в один или несколько разрядов.

Пример коиструкции одиоразрядиого знакового индикатора показан иа рис. 9.4. Структура этого зиакового индикатора сос-

тоит из семи излучающих элементов и децимальной точки, т. е. восьми *p-n*-переходов в одном монокристалле полупроводника, излучающих свет при прохождении тока в прямом направлении. Различные комбинации излучающих элементов, обеспечиваемые внешней коммутацией, позволяют воспроизвести цифры от 0 до 9 и децимальную точку.

Полупроводниковая шкала — это полупроводниковый прибор отображения информации, состоящий из полупроводниковых излучающих элементов, предназначенный для представления аналоговой информации.

Структура полупроводниковой шкалы может представлять собой либо несколько светоизлучающих диодов, расположенных вдоль одной линии, либо несколько *p-n*-переходов, также расположенных по одной линии на общей подложке. Еще одной разновидностью структуры полупроводниковой шкалы является структура с управляемой геометрией светящегося поля (рис. 9.5). Область с электропроводностью *n*-типа низкоомна и поэтому является практически эквипотенциальной. Область с электропроводностью *p*-типа сравнительно высокоомна и поэтому при приложении внешних напряжений к электродам не будет эквипотенциальной. Распределение потенциала в *p*-области зависит от напряжения, поданного на управляющий электрод (рис. 9.5, *б*). Соответственно от поданного на управляющий электрод напряжения зависит и размер светящегося поля полупроводниковой шкалы. Такие полупроводниковые шкалы могут быть использованы в качестве индикаторов настройки транзисторных приемников, для записи аналоговой информации на фотопленку, в качестве шкалы различных измерительных приборов и для других целей.

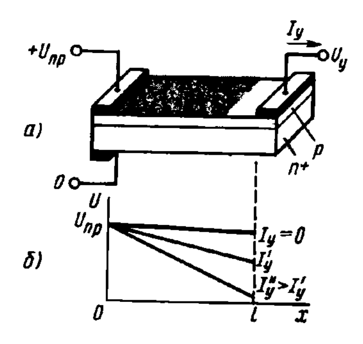


Рис. 9.5. Структура полупроводниковой шкалы с управляемой геометрией светящегося поля (*а*) и распределение потенциала вдоль слаболегированной *p*-области при разных токах через управляющий электрод (*б*)

Полупроводниковый экран — это полупроводниковый прибор отображения информации, состоящий из излучающих элементов, расположенных вдоль одной линии, и содержащий *n* строк излучающих элементов, предназначенный для использования в устройствах отображения аналоговой и цифровой информации.

Примером полупроводникового экрана могут служить полупроводниковые приборы отображения информации АЛ306А... АЛ306И, выпускаемые, однако, как знаковые индикаторы. Они состоят из 36 дискретных светоизлучающих диодов, соединенных в матрицу (семь строк по пять диодов и один диод отдельно) с перекрестной коммутацией и позволяющих воспроизводить цифры и буквы.

Основные характеристики и параметры

Яркость излучения. Яркость излучения является параметром полупроводниковых приборов отображения информации. Единицей яркости в системе СИ служит кандела на квадратный метр (кд/м^2) — яркость источника излучения, каждый квадратный метр излучающей поверхности которого имеет в данном направлении силу света, равную одной канделе. Следует отметить, что световые измерения, строго говоря, не вполне объективны. Основным «прибором», с помощью которого можно измерять светотехнические величины, в конечном счете является глаз человека. Эффективность воздействия света на глаз человека определяется специальной величиной, которая получила название видности. *Видность* V — это отношение светового потока Φ (т. е. оцениваемой нашим глазом мощности) к соответствующей истинной, полной мощности лучистой энергии $\Phi_э$: $V=\Phi/\Phi_э$. На рис. 9.6 приведена зависимость видности от длины волны, определенная Международной комиссией по освещению (МКО). Чувствительность глаза максимальна на длине волны 0,555 мкм. Для стандартного фотометрического наблюдателя 1 Вт лучистой энергии в максимуме чувствительности глаза соответствует 680 лм. Отношение видности света данной длины волны V_λ к максимальной видности V_{max} называют относительной видностью: $K=V_\lambda/V_{max}$.

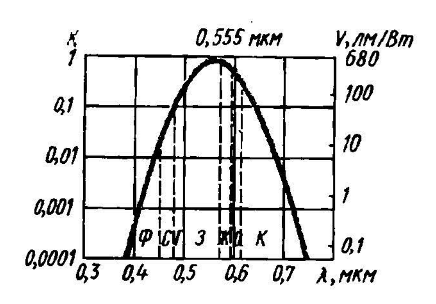


Рис. 9.6. Относительная K и абсолютная V видности стандартного фотометрического наблюдателя в зависимости от длины волны излучения

Таким образом, излучатель, который всю свою энергию отдает только в виде излучения с длиной волны 0,555 мкм, обладает наибольшей яркостью и экономичностью с точки зрения глаза человека. Однако полупроводниковые излучатели часто используют для передачи информации в виде импульсов излучения, которые поступают на приемники излучения со спектральными характеристиками, существенно отличающимися от спектральной характеристики видности глаза человека. В этом случае яркость излучения может оказаться совершенно бесполезным параметром. Так, для инфракрасных излучательных диодов основным параметром является полная мощность излучения в ваттах или милливаттах при определенном прямом токе.

Яркостная характеристика. Полупроводниковые излучатели с выпрямляющим электрическим переходом обладают относительно малым сопротивлением при включении этого перехода в прямом направлении. Поэтому такие излучатели следует считать токовы-

ми приборами, питаемыми от источиков или генераторов тока. Соответственно яркостной характеристикой полупроводниковых приборов отображения информации является зависимость яркости от проходящего через прибор тока. Желательио иметь прямую пропорциональность яркости излучения от проходящего тока, что будет соответствовать неизменности квантового выхода или неизменности отношения излучательных и безызлучательных актов рекомбинации при изменении тока. Аналогом яркостной характеристики для инфракрасных излучающих диодов является зависимость мощности излучения от проходящего тока.

Спектральная характеристика — зависимость мощности излучения от длины волны излучаемых электромагнитных колебаний (рис. 9.7). В первом приближении спектральный состав излучения можно характеризовать цветом свечения полупроводниковых приборов отображения информации, а инфракрасных излучательных диодов — длиной волны излучения в максимуме спектральной характеристики. Но более подробные сведения дает, конечно, спектральная характеристика.

Параметры полупроводниковых излучателей как элементов электрической схемы определяются *вольт-амперной характеристикой*.

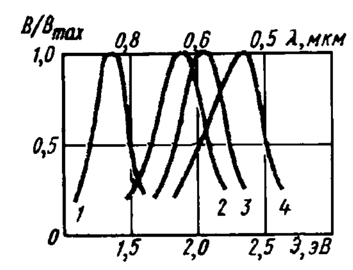


Рис. 9.7. Спектральные характеристики полупроводниковых излучателей на основе арсенида галлия *(1)*, твердого раствора арсенида и фосфида галлия *(2)*, фосфида галлия *(3)* и карбида кремния *(4)*

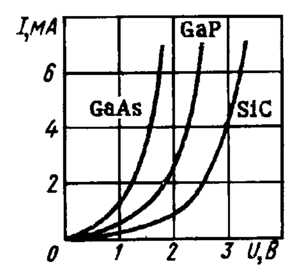


Рис. 9.8. ВАХ полупроводниковых излучателей, изготовленных на основе различных полупроводниковых материалов

Различия прямых ветвей ВАХ полупроводниковых излучателей из разных материалов вызваны прежде всего различием в ширине запрещенной зоны и соответственно в высоте потенциального барьера на *p-n*-переходе (рис. 9.8). Обратные ветви ВАХ не представляют практического интереса, так как полупроводниковые излучатели с выпрямляющим электрическим переходом должны работать только при включении в прямом направлении. Следует, однако, иметь в виду, что пробивные напряжения полупроводниковых излучателей с выпрямляющим электрическим переходом не превышают нескольких вольт.

Инерционность полупроводниковых излучателей характеризу-

ют *временем нарастания импульса излучения* и *временем спада импульса излучения*, которые принято измерять между уровнями излучения 0,1 и 0,9 от амплитуды импульса излучения. Эти времена составляют обычно единицы или десятые доли микросекунды. Таким образом, времена нарастания и спада импульса излучения оказываются несущественными параметрами для полупроводниковых приборов отображения информации, предназначенных для визуальной индикации, так как инерционность глаза человека достаточно велика (около 50 мс). Наоборот, для инфракрасных излучающих диодов, которые предназначены для обработки информации без визуализации, времена нарастания и спада импульса излучения могут быть одними из основных параметров.

§ 9.3. ЭЛЕКТРОЛЮМИНЕСЦЕНТНЫЕ ПОРОШКОВЫЕ ИЗЛУЧАТЕЛИ

Принцип действия

Электролюминесцентные порошковые излучатели — это одна из разновидностей излучающих полупроводниковых приборов, в которых используется электролюминесценция электролюминофора (см. § 1.12). Электролюминесцентный порошковый излучатель представляет собой многослойную систему, состоящую из стеклянной подложки, на которую последовательно наносят проводящий прозрачный электрод из оксидов различных металлов (SnO_2, In_2O_3, CdO и др.), слой электролюминофора, защитный диэлектрический слой в виде лакового покрытия или тонкого слоя диоксида, либо оксида кремния (SiO_2, SiO) и второго электрода (рис. 9.9).

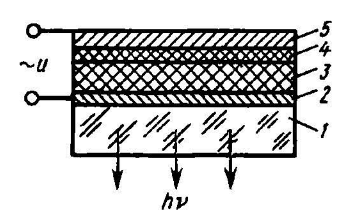


Рис. 9.9. Структура электролюминесцентного порошкового излучателя:

1 — стеклянная подложка; *2* — прозрачный электрод; *3* — слой электролюминофора; *4* — защитный слой; *5* — металлический электрод

Одним из наиболее распространенных электролюминофоров является сульфид цинка ZnS, активированный для получения яркого свечения примесями меди, марганца и других элементов. Зерна или поликристаллы сульфида цинка скрепляются между собой диэлектрической связкой. Так как отдельные зерна порошка сульфида цинка разделены диэлектрической связующей прослойкой, электролюминесцентные конденсаторы могут работать только при переменном напряжении, т. е. возбуждение электролюминофора происходит под действием электрического поля.

Под действием приложенного напряжения происходит ионизация примесных атомов электролюминофора либо в результате туннелирования электронов с примесных уровней в зону проводимости (электростатическая ионизация), либо в результате

ударной ионизации в сильном электрическом поле в обедненных поверхностных слоях зерен сульфида цинка.

После возбуждения поверхностных слоев зерен электролюминофора происходит процесс высвечивания электролюминофора — рекомбинация носителей заряда с выделением избыточной энергии в виде квантов света. Наряду с излучательной рекомбинацией происходит и безызлучательная рекомбинация, при которой избыточная энергия выделяется в виде квантов тепловой энергии.

Цвет излучения определяется шириной запрещенной зоны электролюминофора и глубиной залегания энергетических уровней рекомбинационных ловушек в запрещенной зоне. Длительность процесса высвечивания (послесвечение) зависит от времени жизни неосновных носителей заряда и от наличия в электролюминофоре ловушек захвата, которые могут существенно увеличить эффективное время жизни носителей.

Основные характеристики и параметры

Одним из важнейших параметров электролюминесцентного порошкового излучателя является *эффективная яркость* при определенной частоте переменного напряжения и при определенном значении этого напряжения или плотности тока.

Эффективная яркость электролюминесцентных порошковых излучателей зависит от приложенного напряжения (рис. 9.10). *Яркостная характеристика* нелинейна, так как процесс увеличения избыточной концентрации носителей заряда при ударной ионизации и при туннельном эффекте характеризуется степенными зависимостями от напряжения (или от напряженности электрического поля).

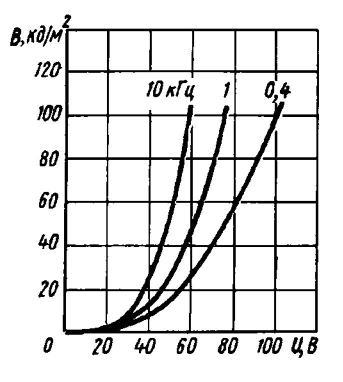


Рис. 9.10. Яркостная характеристика электролюминесцентного порошкового излучателя

Большая нелинейность яркостных характеристик оказывается полезной при создании электролюминесцентных матричных экранов и преобразователей изображения, так как дает возможность получить большую контрастность изображения и большую разрешающую способность.

Крутизну яркостной характеристики иногда оценивают *кратностью изменения яркости* при уменьшении напряжения на электролюминесцентном излучателе в два раза от номинального значения. Кратность изменения яркости электролюминесцентных порошковых излучателей не превышает 25.

Зависимость эффективной яркости от частоты переменного

напряжения (рис. 9.10) объясняется увеличением числа волн яркости в едииицу времени при увеличении частоты.

Спектр излучаемого электролюминесцентным порошковым излучателем света характеризуют длиной волны, соответствующей максимуму спектральной характеристики излучения. Эта длина волны зависит от разности энергий уровней, между которыми происходит переход электронов при излучательной рекомбинации.

§ 9.4. ЭЛЕКТРОЛЮМИНЕСЦЕНТНЫЕ ПЛЕНОЧНЫЕ ИЗЛУЧАТЕЛИ

Принцип действия

Электролюминесцентные пленочные излучатели отличаются от электролюминесцентных порошковых излучателей тем, что между двумя электродами в них находится однородная поликристаллическая пленка электролюминофора, созданная термическим испарением с последующим осаждением в вакууме. Так как в электролюминесцентных пленочных излучателях отсутствует диэлектрическая связка в электролюминофоре, они могут работать и на постоянном токе.

Возбуждение электролюминофора, т. е. создание неравновесного состояния поверхностных слоев отдельных кристаллов электролюминофора, происходит из-за инжекции носнтелей заряда через потенциальные барьеры на поверхности отдельных, контактирующих между собой кристаллов электролюминофора и из-за инжекции из электродов. При рекомбинацни инжектированных носителей избыточная энергия может выделяться в виде квантов света. Возбуждение электролюминофора может происходить также из-за эффектов сильного поля (туннелирование и ударная нонизация) в обедненных поверхностных слоях кристаллов электролюминофора.

Основные характеристики и параметры

Толщина пленки электролюминофора в электролюминесцентных пленочных излучателях мала. Поэтому рабочие напряжения таких излучателей (25...30 В) значительно меньше *рабочих напряжений* электролюминесцентных порошковых излучателей, сделанных из порошкообразного электролюминофора с диэлектрической связкой.

Яркостная характеристика электролюминесцентных пленочных излучателей, т. е. зависимость яркости свечения от приложенного напряжения, нелннейна, так как нелинейны ВАХ этих приборов. В связн с большей нелинейностью яркостных характеристик электролюминесцентных пленочных нзлучателей по сравнению с нелинейностью аналогичных характеристик электролюмннесцентных порошковых излучателей пленочные излучателн обладают большей разрешающей способностью, которая ограни-

чивается размерами отдельных кристаллов электролюминофора ($10^{-5}...10^{-4}$ мм). Разрешающая способность пленочных излучателей велика еще и потому, что тонкие пленки практически не рассеивают свет. *Кратность изменения яркости* электролюминесцентных пленочных излучателей достигает 1000.

Электролюминесцентные пленочные излучатели уступают электролюминесцентным порошковым излучателям по *экономичности и сроку службы*. Низкий срок службы, характерный для большинства полупроводниковых поликристаллических приборов, связан с одновременным действием больших напряженностей электрического поля и повышенной температурой в местах точечных контактов между отдельными кристаллами полупроводника.

Другим недостатком электролюминесцентных пленочиых излучателей, как и электролюминесцентных порошковых излучателей, является большой *разброс параметров*. Этот недостаток также характерен для всех поликристаллических полупроводниковых приборов

§ 9.5. ЛАЗЕРЫ

Прииципп действия

Полупроводниковый лазер — это излучающий полупроводниковый прибор, предназначенный для непосредственного преобразования электрической энергии или энергии некогерентного излучения в энергию когерентного излучения.

В полупроводниковых лазерах или в оптических квантовых генераторах (полупроводниковых ОКГ) излучение, как и в светоизлучающих диодах, порождается рекомбинацией электронов и дырок. Однако эта рекомбинация в лазерах оказывается в основном не самопроизвольной, а вынужденной (стимулированной). Именно поэтому источники вынужденного нзлучения назвали лазерами*. Излучение прн вынужденной рекомбинации получается когерентным (см. § 1.11), что является принципиальным отличием полупроводниковых лазеров от светоизлучающих диодов. Явление вынужденной рекомбинации дает возможность управлять излучением возбужденных атомов полупроводиика с помощью электромагнитных волн и таким образом усиливать и генерировать когерентный свет.

Для работы лазера необходнмо преобладание вынужденной нзлучательной рекомбинации над поглощением квантов света. Преобладание излучения над поглощением или поглощения над излучением зависит от соотношения в кристалле полупроводника возбужденных н невозбужденных атомов, т. е. от населенности энергетических уровней полупроводника. В равновесных условиях на более высоких энергетнческих уровнях при любой температуре полупроводника число электронов меньше, чем на более низких энергетическнх уровнях. При этом нельзя получить усиле-

---

* *Laser — lightwave amplification by stimulated emission of radiation* (усиление световых волн с помощью стимулироваиного излучення).

ния света в результате вынужденной рекомбинации. Для преобладания вынужденной рекомбинации над поглощением квантов света необходимо, чтобы верхние энергетические уровни были больше заполнены электронами, чем нижние. Состояние полупроводника, в котором число электронов на одном из энергетических уровней с большей энергией больше числа электронов на уровне с меньшей энергией, называют *состоянием с инверсной населенностью*. Поглощение квантов света в полупроводнике с инверсной населенностью энергетических уровней мало, так как около потолка валентной зоны почти нет электронов, которым может быть передана энергия кванта света. С другой стороны, в полупроводнике с инверсной населенностью может происходить вынужденная рекомбинация.

Инверсную населенность в полупроводнике можно создать различными способами: 1) с помощью инжекции носителей заряда при прямом включении *p-n*-перехода, что используют в так называемых инжекционных лазерах; 2) путем электронного возбуждения, т. е. путем бомбардировки полупроводника пучком быстрых электронов; 3) с помощью оптической накачки, т. е. путем возбуждения атомов полупроводника квантами света от мощного излучателя некогерентного или когерентного света; 4) путем использования эффектов сильного электрического поля, т. е. лавинного размножения носителей заряда или туннелирования электронов при их переходе с энергетических уровней, расположенных вблизи потолка валентной зоны, на энергетические уровни, расположенные вблизи дна зоны проводимости.

Наибольший практический интерес представляет первый из перечисленных способов создания инверсной населенности. Поэтому рассмотрим инжекционные лазеры.

Конструкция и технология изготовления инжекционных лазеров

Инверсную населенность в инжекционном лазере с *p-n*-переходом проще получить, если одна из областей диодной структуры является вырожденной, т. е. содержит большую концентрацию соответствующих примесей. При прямом включении *p-n*-перехода прямой ток состоит из двух составляющих: электронной и дырочной. Чем больший ток проходит через *p-n*-переход, тем с большим запасом выполняется условие инверсной населенности. Минимальный ток, при котором происходит преимущественно вынужденная рекомбинация, называют *пороговым током*.

Если ток, проходящий через *p-n*-переход, больше порогового, то *p-n*-переход является усиливающей средой для света, распространяющегося в плоскости *p-n*-перехода. Число актов вынужденной рекомбинации можно увеличить, если обеспечить прохождение каждого кванта света несколько раз в плоскости *p-n*-перехода. Для этого две противоположные грани монокристалла

полупроводника делают строго параллельными и тщательно отполированными. Чтобы обеспечить необходимый коэффициент отражения от торцов, можно их не металлизировать, так как большой коэффициент преломления полупроводникового материала обеспечивает отражение от этих торцов до 35% квантов света. После многократного отражения от полированных торцов

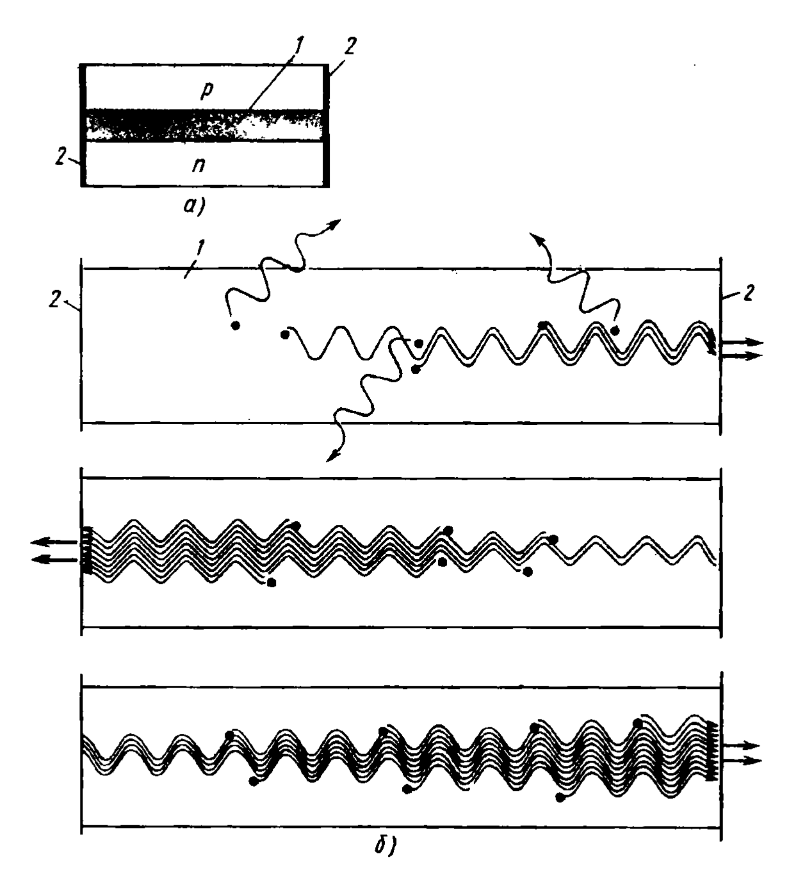


Рис. 9.11. Структура полупроводникового инжекционного лазера *(а)* и схема процесса образования фотонной лавины в оптическом резонаторе лазера *(б)*:

1 — активная область с инверсией населенностью; *2* — отражающие поверхности кристалла полупроводника

и соответствующего многократного прохождения вдоль *p-n*-перехода свет выходит из полупроводника (рис. 9.11). Кванты света, двигающиеся строго перпендикулярно торцам кристалла, могут много раз пройти через активную область с инверсией населенностью и тем самым создать большую лавину квантов света. Две другие боковые грани должны быть скошены под некоторым

углом, чтобы воспрепятствовать возникновению генерации света между ними (рис. 9.12). Те кванты света, которые начали двигаться не вдоль p-n-перехода и не перпендикулярно торцам кристалла, уходят из активной области с инверсной населенностью и не вызывают вынужденной рекомбинации.

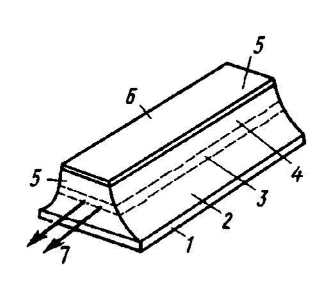


Рис. 9.12. Конструкция полупроводникового инжекционного лазера:

1 — молибденовая пластина — нижний электрод; 2 — область с электропроводностью n-типа; 3 — активная область с инверсной населенностью; 4 — область с электропроводностью p-типа; 5 — полированные торцовые поверхности кристалла полупроводника; 6 — верхний электрод; 7 — излучение

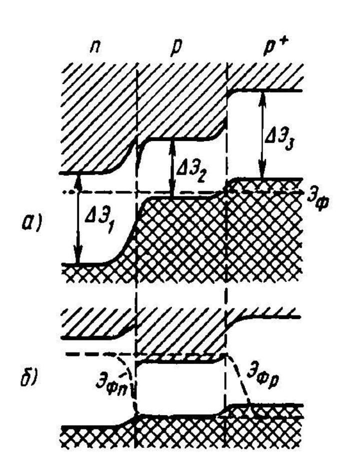


Рис. 9.13. Энергетическая диаграмма полупроводникового инжекционного лазера с гетеропереходами без приложенного напряжения (а) и при прямом напряжении (б)

Для изготовления инжекционных лазеров используют арсенид галлия, твердые растворы арсенид-фосфид галлия $GaAs_{1-x}P_x$, арсенид индия, фосфид индия и другие полупроводниковые материалы. Наиболее широкое распространение получили инжекционные лазеры из арсенида галлия. В качестве исходной заготовки для таких лазеров служит монокристалл арсенида галлия, по форме приближающийся к кубу или параллелепипеду, со сторонами длиной в несколько десятых долей миллиметра.

В арсениде галлия, легированном донорами (Te, Se и др.), p-n-переход обычно создают путем диффузии акцепторов (Zn, Cd и др.). Области с электропроводностью p- и n-типа должны иметь концентрации соответствующих примесей, при которых энергетические состояния электронов и дырок близки к вырождению. Для создания невыпрямляющего контакта с n-областью монокристалл с диодной структурой припаивают к пластинке молибдена, покрытой слоем золота (рис. 9.12). На поверхность p-области наносят слой сплава золота с серебром.

Инверсную населенность значительно проще можно создать в полупроводниковом инжекционном лазере с гетеропереходами (рис. 9.13). Базовую область такой структуры делают из полупроводника с меньшей шириной запрещенной зоны и большей диэлектрической проницаемостью, чем у эмиттерных областей. Инжектированные в базу носители заряда оказываются в потенциальных ямах. Различие показателей преломления (диэлектрических проницаемостей) базовой и эмиттерных областей приводит к полному внутреннему отражению квантов света на гете-

ропереходах, т. е. область базы является, по существу, световодом. Все это обеспечивает значительно меньшие пороговые плотности токов и большие эффективности или КПД инжекционных лазеров с гетеропереходами.

Основные характеристики и параметры

Пороговая плотность тока существенно зависит от температуры инжекционного лазера: для лазеров на основе арсенида галлия пороговая плотность тока порядка 10^2 А/см^2 при $T=4{,}2$ К и порядка 10^4 А/см^2 при 77 К. Таким образом, для уменьшения пороговой плотности тока необходимо глубокое охлаждение инжекционного лазера. Инжекционные лазеры с использованием гетеропереходов, имеющие значительно меньшие пороговые плотности токов, могут работать при комнатной температуре в непрерывном режиме.

Спектральная характеристика лазера, как и любого другого источника света, представляет собой зависимость интенсивности излучения (чаще в относительных единицах) от длины волны (рис. 9.14). При малых токах (меньше порогового) излучение, возникающее в основном из-за самопроизвольной рекомбинации, является некогерентным. Поэтому спектральная характеристика получается широкой, т. е. лазер работает как светоизлучающий диод. При больших токах (больше порогового) интенсивность излучения значительно больше, так как излучение получается когерентным и строго направленным.

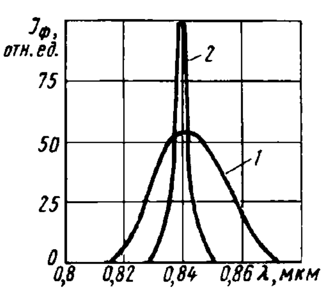


Рис. 9.14. Спектральные характеристики инфракрасного излучающего диода на основе арсенида галлия при температуре 77 К:

1 — при токе ниже порогового значения (режим работы излучающего диода); 2 — при токе выше порогового значения (режим работы инжекционного лазера)

Диаграмма направленности излучения лазера характеризует пространственное распределение интенсивности излучения. Излучение полупроводниковых лазеров обладает достаточно малым (не превышающим нескольких градусов) углом расходимости светового пучка. Но по этому параметру полупроводниковые лазеры значительно уступают газовым и твердотельным диэлектрическим лазерам, что связано с малыми размерами кристалла полупроводника и особенно с малыми размерами активной области, где происходит вынужденная рекомбинация.

Коэффициент полезного действия полупроводникового инжекционного лазера на основе арсенида галлия достигает 70%, в то время как значение внутреннего квантового выхода, возможно, достигает 100%, т. е. каждый инжектированный электрон при рекомбинации с дыркой создает фотон. По КПД полупроводниковые инжекционные лазеры превосходят газовые и твердотельные

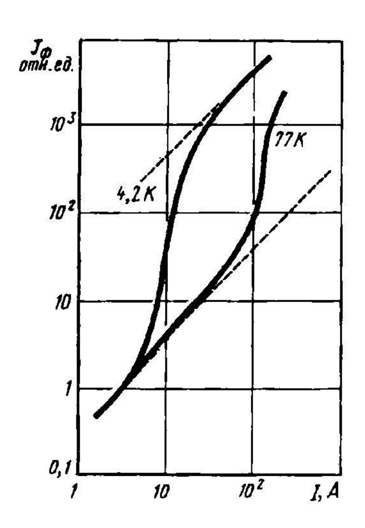


Рис. 9.15. Яркостная характеристика полупроводникового излучающего диода на основе арсенида галлия при различных температурах

диэлектрические лазеры, у которых он равен соответственно тысячным и сотым долям процента.

Яркостная характеристика лазера, т. е. зависимость интенсивности излучения от проходящего через лазер тока (рис. 9.15), представляет собой почти линейные зависимости в диапазоне токов, соответствующих преобладанию самопроизвольной рекомбинации (режим работы светоизлучающего диода) и преобладанию вынужденной рекомбинации (режим работы лазера).

§ 9.6. ФОТОРЕЗИСТОРЫ

Фоторезистор — это полупроводниковый резистор, действие которого основано на фоторезистивном эффекте.

При облучении фоторезистора фотонами в полупроводниковом фоточувствительном слое возникает избыточная концентрация носителей заряда (см. § 1.12). Если к фоторезистору приложено напряжение, то через него будет проходить дополнительная составляющая тока — фототок, обусловленный избыточной концентрацией носителей. Электронная составляющая фототока

$$I_{\phi n} = abq\Delta n\mu_n \frac{U}{l} = abq(1-R)\alpha\eta N_\phi \tau_n \mu_n E\ ,$$

где a — толщина полупроводникового фоточувствительного слоя; b — его ширина; l — расстояние между электродами; R — коэффициент отражения; α — показатель поглощения; η — квантовая эффективность генерации; N_ϕ — число фотонов, падающих на единичную поверхность фоточувствительного слоя в единицу времени.

Фототок соответствует прохождению через фоторезистор и через внешнюю цепь $I_{\phi n}/q$ электронов. Число электронов, возникающих в объеме фоточувствительного слоя из-за поглощения фотонов, равно $abl(1-R)\alpha\eta N_\phi$.

Отношение числа прошедших во внешней цепи электронов к числу возникших в фоточувствительном слое электронов называют *коэффициентом усиления фоторезистора:*

$$K_I = \frac{I_{\phi n}/q}{abl(1-R)\alpha\eta N_\phi} = \frac{ab(1-R)\alpha\eta N_\phi \tau_n \mu_n E}{abl(1-R)\alpha\eta N_\phi} = \frac{\tau_n \mu_n E}{l}\ .$$

Произведение подвижности электронов на напряженность электрического поля есть скорость дрейфа электронов, которую

можно также представить как расстояние между электродами, деленное на время пролета носителей между электродами $t_{прол}$. Поэтому коэффициент усиления фоторезистора можно выразить и в таком виде:

$$K_I = \tau_n / t_{прол} .$$

Если в полупроводниковом фоточувствительном слое есть примеси, являющиеся ловушками захвата для неосновных носителей заряда (сенсибилизирующие или очувствляющие примеси), то захват неосновных носителей этими ловушками может существенно (на несколько порядков) увеличить эффективное время жизни неравновесных основных носителей. В этом случае время жизни может значительно превышать время пролета носителей между электродами. Когда один из электронов достигает положительного электрода, другой электрон входит в полупроводниковый слой из отрицательного электрода для сохранения электрической нейтральности объема полупроводника, в котором осталась нескомпенсированная положительно заряженная ловушка захвата. Таким образом, поглощение одного фотона может служить причиной прохождения через фоторезистор многих электронов.

Введение сенсибилизирующих примесей, приводя к увеличению эффективного времени жизни основных носителей, вызывает снижение быстродействия фоторезистора.

Усиление фототока может происходить и при наличии потенциальных барьеров, например, на поверхности кристаллов полупроводника, если фоторезистор изготовлен на основе поликристаллического полупроводникового материала. Потенциальные барьеры могут являться потенциальными ямами для неосновных носителей заряда. В этом случае будет происходить усиление фототока в фоторезисторе по аналогии с усилением фототока в фототранзисторе, что будет рассмотрено в § 9.9.

Технология изготовления и конструкция

Основной частью конструкции фоторезистора является полупроводниковый фоточувствительный слой, который может быть выполнен в виде монокристаллической или поликристаллической пластинки полупроводника или в виде поликристаллической пленки полупроводника, нанесенной на диэлектрическую подложку. В качестве полупроводникового материала для фоторезисторов обычно используют сульфид кадмия, селенид кадмия или сульфид свинца. На поверхность фоточувствительного слоя наносят металлические электроды. Иногда электроды наносят непосредственно на диэлектрическую подложку перед осаждением полупроводникового слоя.

Поверхность полупроводникового фоточувствительного слоя, расположенную между электродами, называют *рабочей площад-*

кой. Фоторезисторы делают с рабочими площадками прямоугольной формы, в виде меандра или в виде кольца. Площадь рабочих площадок различных фоторезисторов составляет обычно от десятых долей до десятков квадратных миллиметров. Исходя из площади рабочей площадки можно правильно выбрать размер светового пучка, оценить световой поток, при котором должен работать фоторезистор, и т. д. При эксплуатации фоторезистора рекомендуется его рабочую площадку засвечивать полностью, так как при этом эффект изменения сопротивления фоторезистора максимален.

Подложку с нанесенным на нее полупроводниковым фоточувствительным слоем или пластинку полупроводника помещают в пластмассовый или металлический корпус.

Основные характеристики и параметры

Вольт-амперные характеристики фоторезистора представляют собой зависимости светового тока $I_{св}$ при неизменном световом потоке, а также темнового тока $I_{тем}$ от приложенного к фоторезистору напряжения (рис. 9.16). В рабочем диапазоне напряжения ВАХ фоторезисторов при различных значениях светового потока практически линейны. Однако у большинства пленочных фоторезисторов и у фоторезисторов с фоточувствительным слоем из поликристаллического полупроводникового материала линейность ВАХ нарушается при малых напряжениях (характеристика является сверхлинейной). Эта нелинейность связана с явлениями на контактах между отдельными зернами или кристаллами полупроводника. При малых напряжениях сопротивление фоторезистора определяется в основном сопротивлением этих контактов. Напряжение, приложенное к фоторезистору, падает в основном на контактах между зернами полупроводника. Поэтому напряженность электрического поля на контактах получается большой даже при малых напряжениях на фоторезисторе. В связи с этим при увеличении приложенного напряжения сопротивление контактов уменьшается либо из-за эффектов сильного поля (например, туннелирование сквозь тонкие потенциальные барьеры на контактах), либо из-за разогрева приконтактных областей отдельных зерен полупроводника. При дальнейшем увеличении напряжения сопротивление фоторезистора определяется уже объемным сопротивлением зерен полупроводника и поэтому будет оставаться постоянным, что соответствует линейному участку ВАХ.

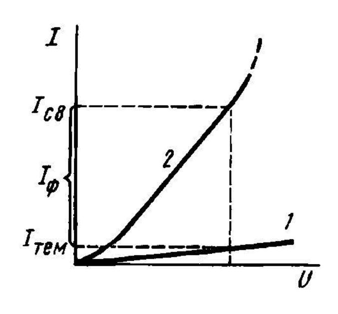


Рис. 9.16. ВАХ фоторезистора:
1 — без облучения (в темноте); 2 — при облучении

При больших напряжениях на фоторезисторе ВАХ опять может отклоняться от линейной, становясь сверхлинейной. Сверхлинейность связана с повышением температуры всего фоточувствительного слоя из-за большой выделяющейся мощности.

Световая, или люкс-амперная, характеристика фоторезистора представляет собой зависимость фототока $I_ф = I_{св} - I_{тем}$ от освещенности, или от падающего на фоторезистор светового потока.

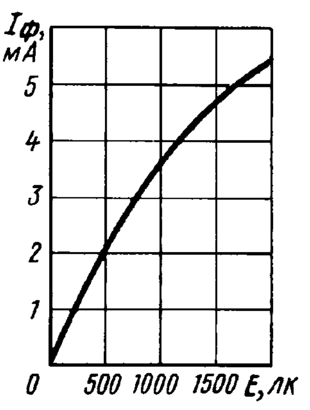


Рис. 9.17. Световая характеристика фоторезистора

Фоторезисторы имеют обычно сублинейную световую характеристику (рис. 9.17). Сублинейность световой характеристики объясняется смещением демаркационных уровней, или квазиуровней Ферми, для электронов и для дырок с увеличением отклонения от равновесного состояния при увеличении освещенности: электронный демаркационный уровень (квазиуровень Ферми для электронов) смещается к зоне проводимости в результате увеличения концентрации свободных электронов, дырочный демаркационный уровень (квазиуровень Ферми для дырок) одновременно смещается к валентной зоне из-за увеличения концентрации дырок (см. рис. 1.8). Вследствие смещения демаркационных уровней часть уровней ловушек захвата становится уровнями рекомбинационных ловушек. С ростом концентрации рекомбинационных ловушек уменьшается время жизни носителей заряда, что и является первой причиной сублинейности световой характеристики. Закономерности возрастания фототока от освещенности различны у различных фоторезисторов и определяются концентрацией тех или иных примесей в полупроводнике и распределением примесных уровней по запрещенной зоне энергетической диаграммы полупроводника.

Второй причиной, приводящей к сублинейности световой характеристики фоторезистора, является уменьшение подвижности носителей заряда при увеличении освещенности из-за увеличения концентрации ионизированных атомов в полупроводнике и, следовательно, из-за увеличения рассеяния носителей заряда ионизированными атомами.

В узком диапазоне освещенностей для аппроксимации световой характеристики часто используют зависимость

$$I_ф = AE^x,$$

где A и x — коэффициенты, являющиеся постоянными для данного фоторезистора в выбранном диапазоне освещенностей; E — освещенность.

Спектральная характеристика фоторезистора — это зависимость фототока от длины волны падающего на фоторезистор света (рис. 9.18). При больших длинах волн, т. е. при малых энергиях квантов света по сравнению с шириной запрещенной зоны полупроводника, энергия кванта оказывается недостаточной для переброса электрона из валентной зоны в зону проводимости. Поэтому для каждого полупроводника и соответственно для каждого фоторезистора существует пороговая длина волны (наибольшая), которую обычно определяют как длину волны, соответствующую спаду фототока на 50% со стороны больших длин волн.

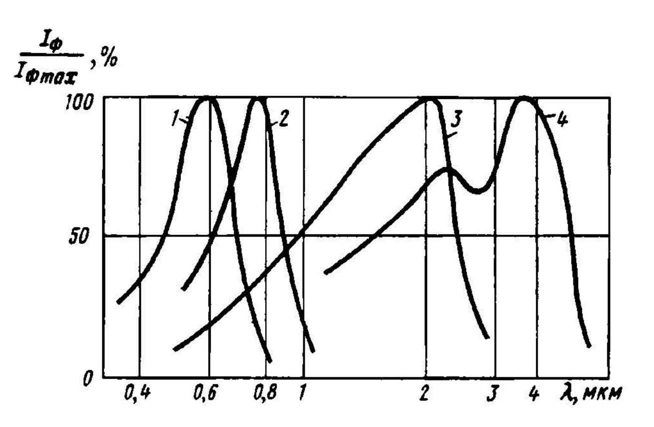


Рис. 9.18. Усредненные спектральные характеристики различных фоторезисторов:
1 — ФСК; 2 — ФСД; 3 — ФСА; 4 — СФ4

При малых длинах волн с уменьшением длины волны падающего на фоторезистор света растет показатель поглощения. Поэтому глубина проникновения квантов света в полупроводник уменьшается, т. е. основная часть неравновесных носителей заряда возникает вблизи освещаемой поверхности фоточувствительного слоя. При этом увеличивается роль поверхностной рекомбинации и уменьшается среднее время жизни неравновесных носителей. Таким образом, спектральная характеристика имеет спад и при малых длинах волн.

Различные полупроводники имеют ширину запрещенной зоны от десятых долей до 3 эВ. Поэтому максимум спектральной характеристики различных фоторезисторов может находиться в инфракрасной, видимой или ультрафиолетовой частях электромагнитного спектра.

Постоянная времени — это время, в течение которого фототок фоторезистора изменяется после освещения или после затемнения фоторезистора на 63% (в *e* раз) по отношению к установив-

шемуся значению. Таким образом, постоянные времени характеризуют скорость реакции фоторезистора на изменение светового потока, т. е. характеризуют инерционность фоторезистора.

В связи с тем, что скорость нарастания фототока при освещении несколько отличается от скорости его спада после затемнения фоторезистора, различают постоянные времени нарастания $\tau_н$ и спада $\tau_{сп}$. Числовые значения постоянных времени различных фоторезисторов от десятков микросекунд до десятков миллисекунд. Постоянные времени измеряют при освещенности 200 лк, температуре окружающей среды 20°C и сопротивлении нагрузки, включенном в схему измерения, меньше 1 кОм. Освещение при определении постоянных времени производят обычно от источника излучения с цветовой температурой 2840 K.

Все эти условия необходимы при измерении постоянных времени для однозначности получаемых результатов, так как постоянные времени зависят от концентрации ловушек захвата и от скорости их заполнения и опустошения, что, в свою очередь, изменяется при изменении освещенности, температуры и других условий, в которых работает фоторезистор. Так, с увеличеннем освещенности уменьшается число ловушек захвата и растет число рекомбинационных ловушек вследствие расщепления уровня Фермн на квазиуровни или смещения демаркационных уровней (см. рис. 1.8). Оба эти фактора приводят к уменьшению времени жизнн носителей заряда и соответственно к уменьшению постоянных времени фоторезистора.

С повышением температуры возрастает вероятность ионизацни ловушек захвата, что означает более быстрое их опустошение и уменьшение постоянных времени.

Наличие существенной инерционности у фоторезисторов приводит к тому, что с увеличением частоты модуляции светового потока эффективное значение возникающего переменного фототока уменьшается. Максимальная частота модуляции светового потока для фоторезисторов не превосходит десятков килогерц.

Темновое сопротивление — это сопротивление фоторезистора в отсутствие освещения. Темновое сопротивление принято определять через 30 с после затемнения фоторезистора, предварительно находившегося под освещенностью 200 лк. Обусловлено это инерционностью опустошения ловушек захвата после прекращения освещения. Например, у фоторезисторов ФСК-1 отношение темновых сопротивлений, измеренных после затемнения через 30 с и через 16 ч, может достигать трех порядков.

Удельная интегральная чувствительность — это отношение фототока к световому потоку и к приложенному напряжению:

$$K_0 = I_ф/(\Phi U).$$

Чувствительность называют интегральной, потому что измеряют ее при освещении фоторезистора светом сложного спектраль-

ного состава: от источника света с цветовой температурой 2840 К при освещенности 200 лк. Удельные интегральные чувствительности различных фоторезисторов составляют от 1 до 600 мА/(В·лм).

§ 9.7. ФОТОДИОДЫ

Полупроводниковый фотодиод — это полупроводниковый диод, обратный ток которого зависит от освещенности.

Обычно в качестве фотодиодов используют полупроводиковые диоды с *p-n*-переходом, смещенным в обратном направлении внешним источником питания. При поглощении квантов света в *p-n*-переходе или в прилегающих к нему областях кристалла полупроводника образуются новые носители заряда (пары электрон — дырка). Неосновные носители, возникшие в прилегающих к *p-n*-переходу областях на расстоянии, не превышающем диффузионной длины, диффундируют к *p-n*-переходу и проходят через него под действием электрического поля или, с точки зрения энергетической диаграммы, скатываются с потенциального барьера (рис. 9.19). Поэтому обратный ток через фотодиод возрастает при освещении. К аналогичному результату приводит поглощение квантов света непосредственно в *p-n*-переходе. В результате при освещении фотодиода обратный ток через него возрастает на величину, называемую фототоком (рис. 9.20).

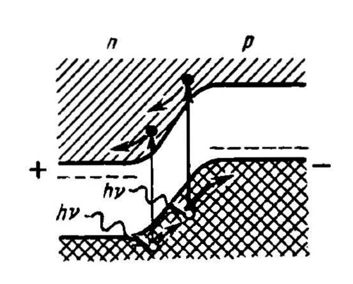


Рис. 9.19. Движение неравновесных носителей заряда, образованных квантами света в *p-n*- переходе или вблизи него, при обратном напряжении на переходе

Рис. 9.20. Обратные ветви ВАХ фотодиода при различных световых потоках

В рабочем диапазоне обратных напряжений при освещении фотодиода обратные токи практически не зависят от приложенного напряжения, хотя обратная ветвь ВАХ фотодиода в затемненном состоянии может не иметь участка насыщения тока.

В конструкции фотодиода, естественно, должна быть предусмотрена необходимость освещения кристалла полупроводника с

одновременной защитой этого кристалла от других внешних воздействий (рис. 9.21).

Свойства фотодиодов можно характеризовать параметрами и зависимостями, аналогичными параметрам и зависимостям фоторезисторов. Однако у фотодиодов есть существенные отличительные особенности. Так, *световая характеристика* фотодиода, т. е. зависимость фототока от освещенности, соответствует прямой пропорциональности фототока от освещенности. Связано это с тем, что толщина базы фотодиода значительно меньше диффузионной длины неосновных носителей заряда. Поэтому практически все неосновные носители, возникшие в базе в результате световой генерации, доходят до *p-n*-перехода и принимают участие в образовании фототока. Во всяком случае потери неосновных носителей заряда на рекомбинацию в базе и на поверхности базы практически не зависят от освещенности, так как исходный полупроводник содержит малое количество неконтролируемых примесей, которые могли бы выполнять роль рекомбинационных ловушек и ловушек захвата.

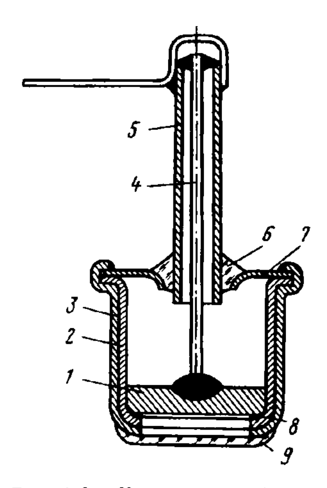
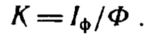


Рис. 9.21. Конструкция фотодиода в металлическом корпусе:

1 — кристалл полупроводника с *p-n*-переходом; 2 — кристаллодержатель; 3 — корпус; 4 — внутренний вывод; 5 — коваровая трубка; 6 — стеклянный проходной изолятор; 7 — ножка корпуса; 8 — кольцо припоя; 9 — стеклянная линза

Следствием линейности световой характеристики фотодиода является независимость интегральной чувствительности фотодиода от приложенного обратного напряжения. Поэтому одним из основных параметров фотодиода является не удельная интегральная чувствительность, а просто *интегральная чувствительность*:



$$K = I_ф / \Phi .$$

Другой особенностью фотодиодов и важным преимуществом их по сравнению с фоторезисторами является малая *инерционность*. Вообще на инерционность фотодиодов могут влиять три физических фактора: время диффузии или дрейфа неравновесных носителей заряда через базу $\tau_д$; время их пролета через *p-n*-переход τ_i; время перезаряда барьерной емкости *p-n*-перехода, характеризующееся постоянной времени $rC_{бар}$.

Время диффузии носителей заряда через базу фотодиода можно определить аналогично времени пролета носителей через базу транзистора по формулам (4.134) или (4.135). В сплавных германиевых фотодиодах толщина базы составляет 20...30 мкм и $\tau_д \approx 50$ нс.

Время пролета носителей через p-n-переход

$$\tau_i \approx \delta / v_{max},$$

где δ — толщина p-n-перехода; v_{max} — максимальная скорость дрейфа носителей заряда (см. § 1.10).

В германии и кремнии $v_{max} \approx 5 \cdot 10^6$ см/с, толщина p-n-перехода, зависящая от обратного напряжения и концентрации примесей в базе, обычно меньше 5 мкм. Следовательно, время пролета носителей через p-n-переход $\tau_i \approx 0{,}1$ нс.

Постоянная времени фотодиода $rC_{бар}$ определяется барьерной емкостью p-n-перехода, зависящей от напряжения, и сопротивлением базы фотодиода при малом сопротивлении нагрузки во внешней цепи. Сопротивление базы у фотодиодов значительно больше, чем у других диодов, так как невыпрямляющий контакт к базе фотодиода расположен по краям базы, а не по всей поверхности (рис. 9.21). Поэтому уменьшение толщины базы может привести не к уменьшению, а к увеличению сопротивления базы. Постоянная времени фотодиодов $rC_{бар}$ получается порядка наносекунд.

Таким образом, инерционность сплавных фотодиодов определяется временем диффузии носителей заряда через базу. В диффузионных фотодиодах, создав ускоряющее электрическое поле в базе из-за неравномерного распределения примесей, можно понизить время пролета носителей через базу до нескольких наносекунд. В таких фотодиодах все три фактора оказывают приблизительно одинаковое влияние на инерционность.

Спектральная характеристика фотодиодов также определяется со стороны больших длин волн шириной запрещенной зоны исходного полупроводникового материала, при малых длинах волн — большим показателем поглощения и увеличением влияния поверхностной рекомбинации носителей заряда с уменьшением длины волны квантов света. Таким образом, коротковолновая граница фоточувствительности фотодиода зависит от толщины базы и от скорости поверхностной рекомбинации. Уменьшая эти величины, можно существенно сдвигать коротковолновую границу фоточувствительности фотодиодов в сторону меньших длин волн.

Положение максимума на спектральной характеристике фотодиода сильно зависит от степени роста коэффициента поглощения в данном полупроводнике. При резком увеличении коэффициента поглощения с уменьшением длины волны падающего света, например в германии, положение максимума определяется шириной запрещенной зоны ($\lambda_{max} = 1{,}55$ мкм) и практически не зависит от толщины базы.

Если зависимость коэффициента поглощения от длины волны слабая, как, например, в кремнии, то эффект уменьшения проникновения квантов света в глубь полупроводника и увеличения роли поверхностной рекомбинации будет сказываться слабее с уменьшением длины волны. Поэтому максимум спектральной

характеристики может смещаться при изменении толщины базы и скорости поверхностной рекомбинации. Так, максимум спектральной характеристики кремниевых фотодиодов можно смещать в диапазоне от 0,6 до 1 мкм путем изменения их конструкции и технологии изготовления.

Фотодиоды на основе перехода металл — полупроводник

Структура фотодиода на основе выпрямляющего перехода металл — полупроводник и его энергетическая диаграмма при обратном напряжении приведены на рис. 9.22. Чтобы основная часть квантов света проникла через металлический верхний электрод *1*, его толщина должна быть малой (около 10 нм для Au). Уменьшения потерь на отражение можно достичь с помощью просветляющих покрытий.

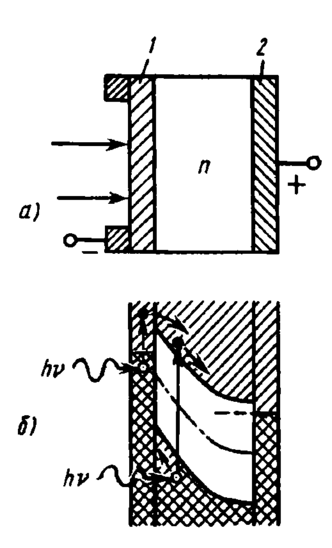


Рис. 9.22. Структура фотодиода с выпрямляющим переходом между металлом и полупроводником (*а*) и энергетическая диаграмма этой структуры при обратном напряжении (*б*):

1 — тонкий прозрачный металлический электрод; *2* — второй металлический электрод, образующий с кристаллом полупроводника омический переход

Принцип действия фотодиода на основе выпрямляющего перехода металл — полупроводник аналогичен принципу действия фотодиода с *p-n*-переходом. Однако есть некоторые различия, которые сказываются на характеристиках и параметрах. Первым отличием является возможность поглощения квантов света с энергией, меньшей ширины запрещенной зоны, для которых полупроводник оказывается прозрачным, в металле верхнего электрода *1*. При этом если энергия кванта света превышает высоту потенциального барьера, то возбужденные электроны из металла могут перейти в полупроводник через потенциальный барьер (рис. 9.22), обеспечив тем самым возникновение фототока. Поэтому длинноволновая граница спектральной характеристики фотодиодов на основе контакта металл — полупроводник определяется высотой потенциального барьера на этом контакте и расположена при более длинных волнах электромагнитного спектра.

Другим отличием рассматриваемых фотодиодов является то, что с уменьшением длины волны квантов света (с ростом энергии квантов) и с увеличением показателя поглощения в полупроводнике кванты света продолжают поглощаться в слое объемного заряда, где существует электрическое поле. Поэтому коротковол-

новая граница спектральной характеристики фотодиодов на основе перехода металл — полупроводник расположена при более коротких волнах электромагнитного спектра.

Таким образом, спектральная характеристика фотодиода на основе перехода металл — полупроводник значительно шире, чем спектральная характеристика фотодиода с *p-n*-переходом из того же полупроводника.

Кроме того, сопротивление базы фотодиода на основе перехода металл — полупроводник намного меньше. Поэтому постоянная времени τ_{rC} оказывается малой и инерционность определяется в основном только временем пролета носителей через область объемного заряда на выпрямляющем переходе металл — полупроводник. Это время пролета может быть порядка 0,1...0,01 нс, что позволяет использовать фотодиоды на основе перехода металл — полупроводник при СВЧ-модуляции светового потока.

Фотодиоды на основе гетероперехода

Энергетическая диаграмма гетероперехода, смещенного в обратном направлении, показана на рис. 9.23. При освещении фотодиода с таким гетеропереходом со стороны широкозонного полупроводника квантами света с энергией $h\nu$ ($\Delta Э_1 > h\nu > \Delta Э_2$) свет поглощается в узкозонном полупроводнике. Широкозонный полупроводник оказывается прозрачным для таких квантов света. Возникающие при этом неосновные носители заряда, проходя через гетеропереход, создают фототок.

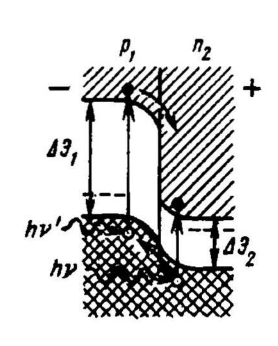


Рис. 9.23. Энергетическая диаграмма гетероперехода при обратном напряжении и при освещении его квантами света с различной энергией ($h\nu' > h\nu$)

С уменьшением длины волны падающего на фотодиод света растет показатель поглощения узкозонного полупроводника. Глубина проникновения квантов в этот полупроводник уменьшается. Генерация неравновесных носителей происходит только вблизи гетероперехода.

При малых длинах волн падающего света ($h\nu' > \Delta Э_1$) кванты света поглощаются в широкозонном полупроводнике. Таким образом, спектральная характеристика фотодиода на основе гетероперехода получается более широкой по сравнению со спектральными характеристиками фотодиодов на основе обычных *p-n*-переходов.

§ 9.8. ПОЛУПРОВОДНИКОВЫЕ ФОТОЭЛЕМЕНТЫ

Принцип действия

Полупроводниковый фотоэлемент — это полупроводниковый прибор с выпрямляющим электрическим переходом, предназначенный для непосредственного преобразования световой энергии в электрическую.

Фотоэлемент работает без внешних источников питания, а сам является источником электрической энергии.

Рассмотрим принцип действия фотоэлемента с *p-n*-переходом в качестве выпрямляющего перехода. При освещении фотоэлемента из-за поглощения квантов света в *p-n*-переходе и областях полупроводника, прилегающих к *p-n*-переходу, происходит генерация новых носителей заряда. Диффузионное электрическое поле, существующее в *p-n*-переходе, производит разделение неравновесных носителей заряда.

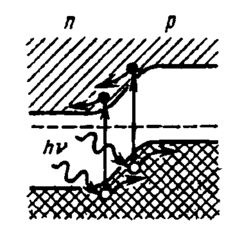


Рис. 9.24. Разделение неравновесных носителей заряда на потенциальном барьере *p-n*-перехода при поглощении квантов света

Другими словами, с точки зрения энергетической диаграммы *p-n*-перехода (рис. 9.24) неравновесные электроны скатываются с потенциального барьера и попадают в *n-область*, дырки, наоборот, в *p*-область. В результате накопления электронов в *n*-области и дырок в *p*-области между этими областями возникает дополнительная разность потенциалов — фото-ЭДС.

Накопление неравновесных носителей в соответствующих областях не может продолжаться беспредельно, так как одновременно с накоплением дырок в *p*-области и электронов в *n*-области происходит понижение высоты потенциального барьера на значение возникшей фото-ЭДС. Уменьшение высоты потенциального барьера или уменьшение суммарной напряженности электрического поля в *p-n*-переходе ухудшает «разделительные свойства» перехода.

Кроме составляющей фото-ЭДС, которая возникает из-за разделения носителей заряда электрическим полем *p-n*-перехода или другого потенциального барьера и которая является основной в фотоэлементах, могут быть и другие составляющие. При освещении полупроводника появляется градиент концентрации электронов и дырок, которые диффундируют от освещаемой поверхности в глубь полупроводника. Но коэффициенты диффузии электронов и дырок различны. Поэтому возникает вторая составляющая фото-ЭДС (см. § 1.12). Кроме того, при наличии на освещаемой поверхности полупроводника ловушек захвата носителей одного знака возникает третья составляющая фото-ЭДС в результате диффузии в глубь полупроводника носителей заряда другого знака.

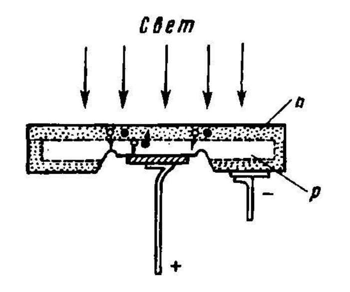


Рис. 9.25. Структура кремниевого фотоэлемента, изготовленного методом диффузии примесей

Технология изготовления и конструкция

В настоящее время фотоэлементы широко применяются в виде солнечных батарей (совокупность электрически соединенных фотоэлементов) для преобразования энергии солнечного света непосредственно в электрическую энергию, питающую установки космических аппаратов. Обычно для этих целей используют кремниевые фотоэлементы. Электронно-дырочный переход в монокристаллической пластинке кремния с электропроводностью *p*-типа создают диффузией фосфора или сурьмы (рис. 9.25). При большой концентрации доноров (фосфор или сурьма) в поверхностном слое кремния проводимость *n*-области получается высокой. Поэтому невыпрямляющий контакт к этой области можно сделать в виде кольца или рамки, оставив всю поверхность кристалла доступной для освещения.

Основные характеристики и параметры

Вольт-амперные характеристики. Режиму работы фотоэлемента (режиму генерации фото-ЭДС) при разных освещенностях или световых потоках соответствуют части ВАХ, расположенные в четвертом квадранте (рис. 9.26). Точки пересечения ВАХ с осью напряжений соответствуют значениям фото-ЭДС или напряжениям холостого хода при разных освещенностях. У кремниевых фотоэлементов фото-ЭДС составляет 0,5...0,55 В.

Точки пересечения ВАХ с осью токов соответствуют значениям токов короткого замыкания, которые зависят от площади выпрямляющего электрического перехода фотоэлемента. Поэтому сравнивают и оценивают фотоэлементы по плотностям тока короткого замыкания. У кремниевых фотоэлементов плотность тока короткого замыкания при средней освещенности солнечным светом составляет 20..25 мА/см^2.

Рис. 9.26. ВАХ фотоэлемента при различных световых потоках, падающих на фотоэлемент

По ВАХ при разных освещенностях фотоэлемента можно выбрать оптимальный режим работы фотоэлемента, т. е. оптимальное сопротивление нагрузки, при котором в нагрузке выделяется наибольшая мощность. Оптимальному режиму работы фотоэлемента соответствует наибольшая площадь

прямоугольиика с вершиной на ВАХ при даииой освещениости (рис. 9.26).

Для кремииевых фотоэлементов при оптимальиой иагрузке напряжеиие иа иагрузке 0,35...0,4 В, плотиость тока через фотоэлемент 15...20 мА/см^2.

Световые характеристики фотоэлемента — это зависимости фото-ЭДС и тока короткого замыкания от светового потока или от освещеииости фотоэлемеита (рис. 9.27). Сублииейность световых характеристик связаиа с уменьшением высоты потеициального барьера при иакоплении избыточного заряда электроиов в n-области и дырок в p-области.

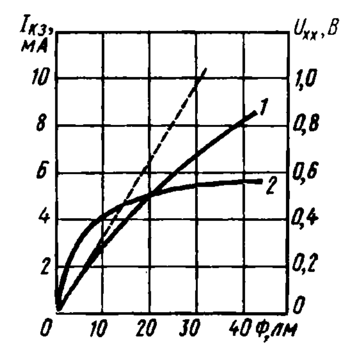


Рис. 9.27. Световые характеристики фотоэлемента:
1 — при коротком замыканни; 2 — при холостом ходе

Спектральная характеристика фотоэлемента — это зависимость тока короткого замыкаиия от длииы волны падающего света. Спектральные характеристики фотоэлемеитов аиалогичны спектральным характеристикам фотодиодов, изготовлениых на основе того же полупроводиика. Максимум спектральиой характеристики кремниевых фотоэлементов почти соответствует максимуму спектрального распределеиия эиергии солнечного света. Имеиио поэтому кремниевые фотоэлементы широко используют для создания солнечных батарей.

Коэффициент полезного действия фотоэлемента — это отиошеиие максимальной мощиости, которую можио получить от фотоэлемента, к полной мощности лучистого потока, падающего на рабочую поверхность фотоэлемента:

$$\eta = P_{max}/P.$$

К основным процессам, приводящим к умеиьшению КПД фотоэлемеитов, относят отражение части излучения от поверхиости полупроводиика, фотоэлектрически иеактивиое поглощеиие кваитов света в полупроводиике (без образоваиия пар иосителей электрон — дырка), рекомбинацию иеравиовесиых носителей еще до их разделения электрическим полем выпрямляющего электрического перехода (особенно на поверхности кристалла полупроводника), потери мощности при прохождеиии тока через объемное сопротивление базы фотоэлемеита.

В результате при преобразоваиии солиечного света в электрическую энергию КПД кремииевых фотоэлементов ие превышает 12%. Однако его можно существенно повысить, используя в качестве исходного полупроводника теллурид кадмия, арсеиид галлия и другие материалы с несколько большей шириией запрещенной зоны, чем у кремния, а также используя фотоэлемеиты на основе гетеропереходов.

§ 9.9. ФОТОТРАНЗИСТОРЫ И ФОТОТИРИСТОРЫ

Транзистор, реагирующий на облучение световым потоком и способный одновременио усиливать фототок, называют фототранзистором.

Биполярные фототранзисторы

Биполярный фототранзистор может быть включен в схему по-разному. Если подать напряжеиие между базой и коллектором, сместив коллекторный переход в обратном иаправлеиии и оставив эмиттерный вывод неподключенным к схеме, то такое включение бнполярного фототранзистора иичем не будет отличаться от схемы включения фотодиода. При поглощеиии кваитов света в базовой и коллекторной областях образуются неравновесные пары иосителей заряда (электроны и дырки). Неосиовные носители (дырки в n-базе и электроны в p-коллекторе для транзистора p-n-p-типа) диффундируют к коллекторному переходу, втягиваются существующим там электрическим полем в коллекторный переход и проходят через него, создавая тем самым фототок $I_ф$.

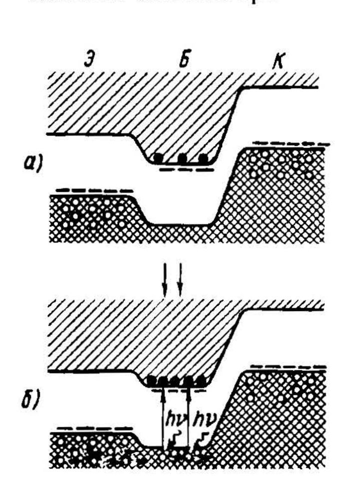


Рис. 9.28. Эиергетическая диаграмма фототранзистора, находящегося в затемиеииом состоянии (а) и при освещении (б)

Однако биполяриый фототранзистор обычно применяют при включении его по схеме с общим эмиттером. Поэтому рассмотрим принцип действия биполярного фототранзистора, включеиного по схеме с общим эмиттером. Предположим виачале, что базовый вывод ие подключен к схеме, т. е. ток базы равен нулю ($I_Б = 0$). В этом случае неосновиые носители заряда, проходя через p-n-переход коллектора, создают тот же фототок $I_ф$. Неравновесиые осиовные носители — электроны в n-базе, возникшие из-за поглощеиия там квантов света, и электроны, пришедшие в базу из коллектора, — оказываются в своеобразной потенциальной яме (рис. 9.28). Накопление в базе иеравиовесных основиых носителей заряда поиижает высоту потенциальных барьеров эмиттериого и коллекториого переходов. Из-за уменьшения высоты потенциального барьера эмиттерного перехода увеличивается инжекция дырок из эмиттера в базу (рис. 9.28, б). Соответственно возрастает и ток коллектора. Таким образом, иакопленный в базе биполярного фототранзистора дополнительный заряд неравновесных основных носителей обеспечивает усиление фототока, т. е. при освещении результирующий ток коллектора $I'_ф \approx h_{21э} I_ф$.

При подключении вывода базы к внешней схеме ток базы может измеияться при освещении фототранзистора. Степепь изменения этого тока завнсит от сопротивлеиий в цепи базы. Измеиеиие тока базы происходит в результате выхода иеравиовесных электронов из иее во виешиюю базовую цепь. В результате иакоплеиный в базе заряд осиовных иосителей умеиьшается, что умеиьшает усиление фототока.

Таким образом, биполяриый фототраизистор обладает наибольшей чувствительностью к облучеиию светом базовой области при включеиии по схеме с общим эмиттером и отключеииой базе. Поэтому у первых конструкций биполяриых фототраизисторов вывод базы вообще отсутствовал. Одиако иаличие вывода базы у биполярных фототранзисторов позволяет использовать ие только оптическое, но и электрическое управлеиие фототраизистором, осуществлять компенсацию посторонних внешиих воздействий (иапример, измеиеиие параметров, вызванное изменеиием температуры в процессе работы).

Фототиристоры

Фототиристором называют тиристор, напряжение включения которого уменьшается с увеличением освещенности.

Соответственно отличительиой особенностью структуры и коиструкции фототиристора является возможность освещеиия одиой из базовых областей (рис. 9.29).

Переключение фототиристора из закрытого в открытое состояиие происходит, как и у обычного тиристора, при увеличении суммарного коэффициента передачи по току тиристорной структуры до единицы (см. § 5.1). В фототиристоре увеличеиие этого параметра может происходить в результате увеличения тока через тиристорную структуру при поглощении квантов света в базовых областях, т. е. из-за генерации носителей заряда в базовых областях при их освещении. Таким образом, фототиристор является аиалогом управляемого тиристора, включение которого в открытое состояние может быть произведеио импульсом света.

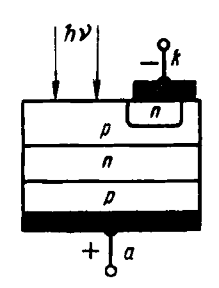


Рис. 9.29. Структура фототиристора

§ 9.10. ПРИЕМНИКИ ПРОНИКАЮЩЕЙ РАДИАЦИИ И КОРПУСКУЛЯРНО-ПРЕОБРАЗОВАТЕЛЬНЫЕ ПРИБОРЫ

Взаимодействие частиц высокой энергии и проникающей радиации с полупроводником

Частицы с высокой энергией и проникающая радиация, проходя через полупроводник, могут производить ионизацию атомов полупроводника. Каждая частица с высокой энергией или каждый квант проникающей радиации могут образовать в полупроводнике большое число свободных электронно-дырочных пар — носителей заряда. Этот процесс является основой принципа действия различных полупроводниковых приемников проникающей радиации и корпускулярно-преобразовательных приборов (датчиков проникающей радиации без *p-n*-перехода и с *p-n*-переходом, атомных электроэлементов).

Однако при воздействии проникающей радиации или частиц высокой энергии кроме образования свободных носителей заряда в полупроводнике могут происходить и нежелательные процессы. Так, атомы кристаллической решетки могут быть смещены из положения равновесия с образованием различных дефектов кристаллической решетки. Например, облучение германия электронами высокой энергии аналогично введению в кристаллическую решетку германия атомов акцепторной примеси — в запрещенной зоне полупроводника образуются акцепторные примесные уровни.

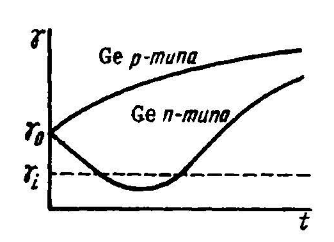


Рис. 9.30. Зависимость удельной проводимости германия от времени облучения его электронами больших энергий

После облучения германия *n*-типа электронами с энергией около 1 МэВ проводимость уменьшается. Это связано с уменьшением концентрации носителей заряда в связи с компенсацией. Когда концентрация образованных облучением акцепторных дефектов равна концентрации основных носителей заряда (электронов), наблюдается минимум проводимости. Дальнейшее облучение обусловливает возрастание концентрации дырок и увеличение проводимости (рис. 9.30). Необходимо отметить, что возникновение дополнительных акцепторных примесей в кристалле приводит к уменьшению подвижности носителей. Однако относительное уменьшение проводимости из-за уменьшения подвижности значительно меньше увеличения проводимости из-за изменения концентрации дырок. После облучения германия *p*-типа электронами высокой энергии его проводимость возрастает.

При облучении полупроводника медленными нейтронами в нем происходят радиоактивные изменения, в результате которых в исходном кристалле появляются примеси посторонних элементов. В германии, например, при нейтронном облучении появляются примеси галлия и мышьяка. Так как образование атомов галлия при этом происходит в три раза быстрее, чем атомов мышьяка, то общая электропроводность определяется избытком галлия, т. е. бомбардировка германия нейтронами приводит к появлению дырочной электропроводности.

При облучении полупроводника быстрыми нейтронами наряду с радиоактивными превращениями в кристалле происходит образование дефектов, подобных тем, какие образуются под действием быстрых электронов, и приводящих к появлению акцепторных уровней в кристалле.

Приемники проникающей радиации на основе поликристаллических полупроводников

Как принцип действия, так и метод изготовления приемников проникающей радиации на основе поликристаллических полупроводников аналогичен методу изготовления поликристаллических фоторезисторов. Основой приемника проникающей радиации является поликристаллический слой селенистого или сернистого кадмия, нанесенный на проводящую подложку путем, например, возгонки в вакууме. Эта подложка служит одновременно одним из электродов приемника проникающей радиации. Второй электрод наносят путем термического испарения в вакууме почти на всю поверхность поликристаллического полупроводникового слоя. Этот второй электрод в связи с малой его толщиной является прозрачным для рентгеновского излучения и тем более для гамма-излучения. Таким образом, в отличие от фоторезисторов в приемниках проникающей радиации электроды расположены на противоположных поверхностях поликристаллического полупроводникового слоя.

Отечественная промышленность выпускает приемники проникающей радиации различных марок: РГД-0, РГД-1, РГД-2 и ГД-Г1 (рентгено-гамма-датчик и гамма-датчик). Отличаются эти приемники проникающей радиации друг от друга в основном внешним конструктивным оформлением. Чувствительностью к рентгено- и гамма-излучению помимо перечисленных приемников проникающей радиации обладают сернисто-кадмиевые фоторезисторы ФСК-М и ФСК, а также селенисто-кадмиевые ФСД. Однако удельная чувствительность фоторезисторов к проникающей радиации примерно на два порядка меньше, чем приемников проникающей радиации. Свойства приемников проникающей радиации характеризуются зависимостями и параметрами, аналогичными зависимостям и параметрам фоторезисторов. Основ-

ным недостатком приемников проникающей радиации является их большая инерционность, особенно при малых интенсивностях облучения. Постоянные времени приемников проникающей радиации составляют несколько десятков секунд при мощности облучения 5 Р/ч.

Приемники проникающей радиации и частиц высокой энергии на основе монокристаллов

Принцип действия такого приемника проникающей радиации и частиц высокой энергии аналогичен принципу действия фотодиода. Приемник проникающей радиации или частиц высокой энергии на основе монокристаллов полупроводника с *p-n*-переходом используют при подаче на него обратного напряжения. Поглощение в полупроводнике частиц высокой энергии или проникающей радиации приводит к возникновению новых неосновных носителей заряда и, следовательно, к увеличению обратного тока.

Рабочим объемом такого приемника в основном является объем *p-n*-перехода, т. е. область, обедненная подвижными носителями заряда. При проникновении в кристалл полупроводника частицы высокой энергии вдоль ее трека образуется колонка ионизации, геометрические размеры которой могут быть относительно большими, так как пробег частицы высокой энергии в полупроводнике может составлять десятки сантиметров. Поэтому толщина *p-n*-перехода приемника проникающей радиации или частиц высокой энергии должна существенно превышать толщину *p-n*-перехода фотодиода и других диодов.

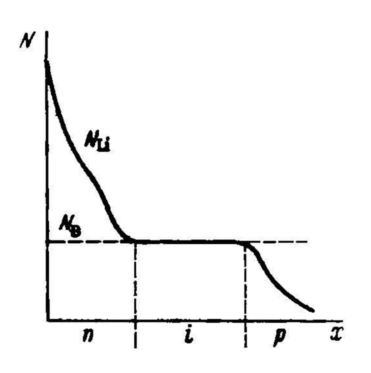


Рис. 9.31. Распределение концентрации примесей бора и лития по кристаллу кремния после диффузии и дрейфа лития

Наиболее существенные результаты в создании диодов с большой толщиной *p-n*-перехода достигнуты при использовании диффузии атомов лития в кремний или германий с электропроводностью *p*-типа с последующим дрейфом ионов лития. Литий, представляющий собой донорную примесь для кремния и германия, имеет чрезвычайно высокий коэффициент диффузии, а ионы лития — большую подвижность при наличии электрического поля в монокристаллах кремния или германия. При изготовлении, например, кремниевых приемников проникающей радиации используют толстые (более 10 мм) монокристаллические пластинки высокоомного кремния с примесью бора. Диффузия лития с одной стороны пластинки кремния проводится при относительно низкой температуре

(300...600°С) в течение нескольких минут. В дальнейшем к полученному *p-n*-переходу прикладывают обратное напряжение при одновременном нагреве кристалла до температуры около 200°С. Из-за большой подвижности ионы лития дрейфуют в электрическом поле, компенсируя исходную проводимость или заряды ионов бора в *p-n*-переходе. В результате получается распределение примесей по толщине кристалла, показанное на рис. 9.31. Таким методом создают диоды с толщиной *p-n*-перехода до 10 мм.

Приемники проникающей радиации на основе монокристаллов полупроводников с *p-n*-переходом, имеющим большую толщину, позволяют регистрировать и спектрометрировать длиннопробежные частицы высокой энергии. Основным недостатком таких приемников, созданных с использованием диффузии и дрейфа лития, является необходимость их глубокого охлаждения при хранении и транспортировке во избежание самопроизвольной диффузии и самопроизвольного дрейфа лития, т. е. во избежание перераспределения примесей.

Атомные электроэлементы

С генерацией ЭДС при облучении *p-n*-перехода частицами высокой энергии связан один из возможных путей создания атомного электроэлемента, преобразующего энергию радиоактивных излучений в электрическую энергию.

Первые атомные электроэлементы работали с радиоактивным стронцием ^{90}Sr, представляющим собой «золу» атомных реакторов. Стронций при распаде испускает электроны больших энергий, которые проникают в толщу полупроводниковой пластинки с *p-n*-переходом и производят там ионизацию атомов полупроводника. Возникшие пáры носителей разделяются диффузионным полем перехода, что обеспечивает появление нескомпенсированных зарядов на электродах атомного электроэлемента. В результате в первых атомных электроэлементах создавалась разность потенциалов до 0,2 В. Ток короткого замыкания атомного электроэлемента площадью 0,32 $см^2$ составлял 5 мкА, коэффициент полезного действия — 1%. Мощность, отдаваемая одним атомным электроэлементом, невелика, поэтому их соединяли последовательно и параллельно, создавая таким образом атомные батареи. Поскольку период полураспада стронция равен 28 годам, предполагалось, что срок службы атомных батарей должен исчисляться по крайней мере 28 годами. Однако оказалось, что эффективность атомных батарей значительно снижалась в течение двух-трех недель из-за того, что быстрые электроны, излучаемые стронцием, создавали дефекты в кристаллической решетке полупроводника, портили структуру *p-n*-перехода.

Дальнейшая разработка атомных батарей в основном ведется по трем направлениям: 1) подбирают радиоактивные вещества, спектр излучения которых не содержит очень быстрых частиц;

2) предполагают использовать замедлители частиц, с тем чтобы *p-n*-переход подвергался облучению, но не выходил из строя; 3) энергию испускаемых радиоактивным веществом частиц пытаются предварительно преобразовать в световую. Для этого на пути движения частиц высокой энергии помещают слой люминофора. Поглощаемые люминофором частицы вызывают его свечение, которое воздействует на *p-n*-переход фотоэлемента.

§ 9.11. ОПТОПАРЫ И ОПТОЭЛЕКТРОННЫЕ МИКРОСХЕМЫ

Оптопары

Оптопара — это оптоэлектронный полупроводниковый прибор, состоящий из излучающего и фотоприемного элементов, между которыми имеется оптическая связь и обеспечена электрическая изоляция.

В качестве одного элемента оптопары — излучателя — могут быть использованы инфракрасный излучающий диод, светоизлучающий диод, электролюминесцентный порошковый или пленочный излучатель, а также полупроводниковый лазер. Обычно применяют инфракрасный излучающий диод, который отличается от светоизлучающего диода и от электролюминесцентных порошковых и пленочных излучателей большим КПД (превышающим 10%), а также простотой структуры и управления.

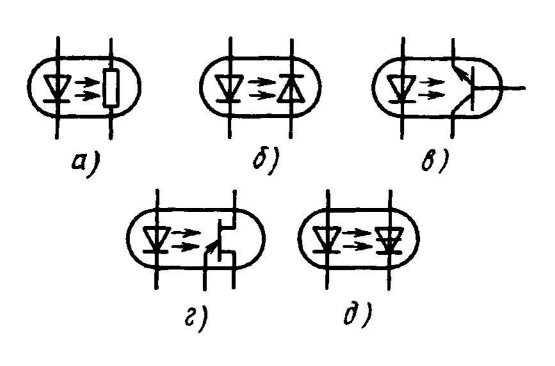


Рис. 9.32. Оптопары:
а — резисторная; *б* — диодная; *в* — транзисторная с биполярным фототранзистором; *г* — транзисторная с однопереходным фототранзистором; *д* — тиристорная

В качестве второго элемента оптопары — фотоприемного элемента — могут быть использованы фоторезистор, фотодиод, биполярный фототранзистор, однопереходный фототранзистор или фототиристор. В зависимости от вида фотоприемного элемента различают резисторные, диодные, транзисторные и тиристорные оптопары. Излучающий и фотоприемный элементы оптопары помещают в общий корпус. Примеры схемного обозначения различных оптопар приведены на рис. 9.32.

Усиление или преобразование входного электрического сигнала в оптопарах происходит следующим образом. Изменение входного тока — тока через полупроводниковый излучатель — сопровождается изменением мощности его излучения и изменением

потока фотонов, падающих на фотоприемный элемент оптопары. При изменении облучения фотоприемного элемента происходит либо изменение сопротивления фоторезистора, либо изменение обратного тока фотодиода, либо появление фото-ЭДС при работе фотодиода в режиме фотоэлемента, либо усиление фототока в фототранзисторе, либо переключение из закрытого состояния в открытое фототиристора или однопереходного фототранзистора. Следует отметить универсальность однопереходного фототранзистора в качестве фотоприемного элемента оптопар. Он может быть использован на выходе оптопары как фоторезистор, фотодиод, фотоэлемент и, конечно, как переключающийся прибор — однопереходный фототранзистор.

Для существования хорошей оптической связи между элементами оптопары кроме их соответствующего расположения необходимо по возможности более близкое совпадение спектральных характеристик этих элементов.

Оптоэлектронные интегральные микросхемы

Обязательной частью любой оптоэлектронной интегральной микросхемы является та или иная оптопара. В зависимости от необходимого быстродействия обработки аналоговых или логических сигналов, от необходимой мощности на выходе и от других требований в качестве фотоприемного элемента оптопары применяют тот или иной из перечисленных в этом параграфе элементов. Каждый из них имеет свои преимущества и свои недостатки. Так, фотодиод может обладать большим быстродействием, но его коэффициент усиления фототока не превышает единицы. Поэтому всегда между выходом оптопары и исполняющим устройством есть согласующая электронная схема. При современном уровне интегральной технологии введение такой согласующей схемы в интегральном исполнении, объединенной с оптопарой в едином корпусе, не представляет принципиальной сложности. Но благодаря наличию оптической связи оптоэлектронные интегральные микросхемы, как и простые оптопары, обладают рядом существенных достоинств.

1. Почти идеальная гальваническая развязка управляющих цепей от исполнительных при сохранении сильной функциональной связи (оптической) между ними. В отличие от электрической оптическая связь может быть легко реализована между низкоомными и высокоомными, низковольтными и высоковольтными, низкочастотными и высокочастотными цепями. Развязка электрических цепей с помощью оптопар достигает 10^{16} Ом и 10^5 В, а емкость между цепями может составлять всего 10^{-4} пФ, причем эти параметры могут быть улучшены. Здесь же можно отметить однонаправленность оптической связи и отсутствие обратной реакции приемника излучения на излучатель.

2. Помехозащищенность оптических каналов связи, что является следствием зарядовой нейтральности фотонов. Поток фотонов в тракте передачи сигнала не вызывает появления индуцированных электрических и магнитных полей, что ограничивает плотность электронного потока и является причиной различных паразитных связей в электронных цепях. Высокая помехозащищенность оптического канала связи и отсутствие взаимных наводок позволяют выполнять многоканальные сложные связи.

3. Для передачи информации по оптическому каналу связи могут быть использованы изменения направления, интенсивности, спектрального состава, поляризации или фазы колебаний. Это открывает возможности для параллельной обработки информации.

4. Возможность выдачи выходной информации, а также информации о состоянии отдельных звеньев оптоэлектронного устройства в виде световых сигналов или зрительно воспринимаемого образа, что позволяет осуществлять непосредственный, визуальный контроль и считывание.

Контрольные вопросы

1. Каким образом в полупроводниковом излучателе происходит непосредственное преобразование электрической энергии в энергию электромагнитного излучения?

2. Как должны быть выполнены структура и конструкция светоизлучающего диода для получения наибольшего внешнего квантового выхода?

3. Какими параметрами можно характеризовать различные свойства светоизлучающих диодов?

4. Каков принцип действия полупроводникового лазера?

5. Каковы отличия в принципе действия и в свойствах полупроводникового лазера и светоизлучающего диода?

6. Почему когерентное излучение от инжекционного полупроводникового лазера можно получить только при токах, превышающих некоторое пороговое значение?

7. Как объяснить спектральную характеристику фоторезистора?

8. Что такое коэффициент усиления фоторезистора и как это усиление происходит?

9. Какими параметрами характеризуют фоторезистор?

10. Какие физические факторы влияют на световую характеристику фоторезистора при больших световых потоках?

11. Каковы отличия в свойствах фотодиода и фоторезистора?

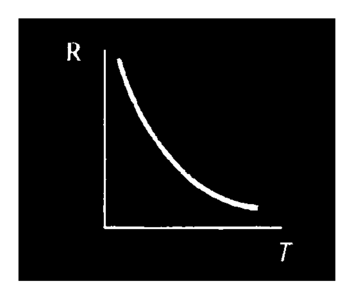
12. Какие структуры могут иметь фотодиоды и каковы основные отличия в свойствах фотодиодов на основе различных выпрямляющих электрических переходов?

13. Как в фотоэлементе происходит непосредственное преобразование световой энергии в электрическую?

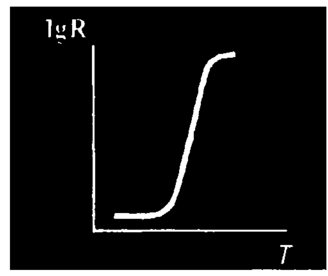
14. Каковы отличия в принципе действия и в свойствах фотодиода и биполярного фототранзистора?

15. Почему фототиристор может управлять относительно большими мощностями, чем допустимая мощность рассеяния самого фототиристора?

16. Перечислите достоинства оптопар и оптоэлектронных интегральных микросхем.



10



Терморезисторы

Терморезистор — это резистор, в котором используется зависимость электрического сопротивления полупроводника от температуры.

§ 10.1. ТЕРМИСТОРЫ ПРЯМОГО ПОДОГРЕВА

Термистор — это полупроводниковый терморезистор с отрицательным температурным коэффициентом сопротивления.

Принцип действия термисторов

В *термисторах прямого подогрева* сопротивление изменяется или под влиянием теплоты, выделяющейся в них при прохождении электрического тока, или в результате изменения температуры термистора из-за изменения его теплового облучения (например, при изменении температуры окружающей среды).

Уменьшение сопротивления полупроводника с увеличением температуры (отрицательный температурный коэффициент сопротивления) может быть вызвано разными причинами — увеличением концентрации носителей заряда, увеличением интенсивности обмена электронами между ионами с переменной валентностью или фазовыми превращениями полупроводникового материала.

1. Первое явление характерно для термисторов, изготовленных из монокристаллов ковалентных полупроводников (кремний, германий, карбид кремния, соединения типа $A^{III}B^{V}$ и др.). Такие полупроводники обладают отрицательным температурным коэффициентом сопротивления в диапазоне температур, соответствую-

щих примесной электропроводности, когда не все примеси ионизированы, а также в диапазоне температур собственной электропроводности, когда концентрация носителей изменяется из-за ионизации собственных атомов полупроводника (см. рис. 1.9). И в том и в другом случае зависимость сопротивления полупроводника определяется в основном изменением концентрации носителей заряда, так как температурные изменения подвижности при этом пренебрежимо малы.

В этих диапазонах температур зависимость сопротивления полупроводника от температуры соответствует уравнению

$$R = R_\infty \exp \frac{B}{T}, \qquad (10.1)$$

где B — коэффициент температурной чувствительности; R_∞ — коэффициент, зависящий от материала и размеров термистора.

При неполной ионизации примесей и отсутствии компенсации

$$B \approx \Delta Э_п / (2k),$$

где $\Delta Э_п$ — энергия ионизации примесей (доноров или акцепторов).

Для скомпенсированного полупроводника при неполной ионизации примесей

$$B \approx \Delta Э_п / k. \qquad (10.2)$$

При собственной электропроводности

$$B \approx \Delta Э / (2k), \qquad (10.3)$$

где $\Delta Э$ — ширина запрещенной зоны полупроводника.

2. Основная часть термисторов, выпускаемых промышленностью, изготовлена из оксидных полупроводников — оксидов металлов переходной группы таблицы Д. И. Менделеева (от титана до цинка). Такие термисторы в форме стержней, трубок, дисков или пластинок получают методом керамической технологии, т. е. путем обжига заготовок при высокой температуре.

Электропроводность оксидных полупроводников с преобладающей ионной связью отличается от электропроводности ковалентных полупроводников. Для металлов переходной группы характерны наличие незаполненных электронных оболочек и переменная валентность. В результате при образовании оксида в определенных условиях (наличие примесей, отклонение от стехиометрии) в одинаковых кристаллографических положениях оказываются ионы с разными зарядами. Электропроводность таких материалов связана с обменом электронами между соседними ионами. Энергия, необходимая для такого обмена, экспоненциально уменьшается с увеличением температуры. В результате изменения интенсивности обмена электронами между ионами температурная зависимость сопротивления термистора из оксидного полупроводника имеет такой же характер, как и у термисторов из ковалентных полупроводников (рис. 10.1), но коэффи-

циент температурной чувствительности в этом случае отражает изменение интенсивности обмена электронами между ионами, а не изменение концентрации носителей заряда.

3. В оксидах ванадия V_2O_4 и V_2O_3 при температуре фазовых превращений (68 и —110°С) наблюдается уменьшение удельного сопротивления на несколько порядков. Это явление также может

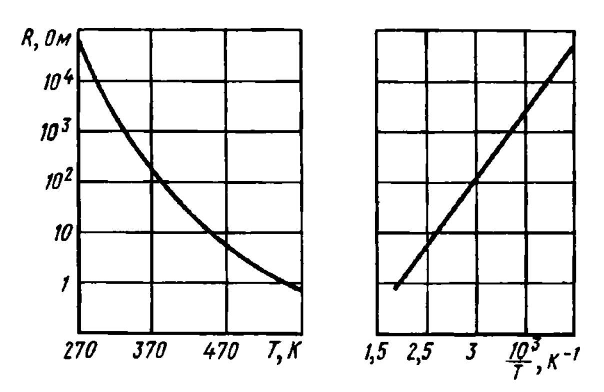


Рис. 10.1. Температурная характеристика одного из термисторов

быть использовано для создания термисторов с большим отрицательным температурным коэффициентом сопротивления в диапазоне температур, соответствующих фазовому превращению.

Характеристики и параметры термисторов прямого подогрева

Температурная характеристика термистора — это зависимость его сопротивления от температуры. Пример температурной характеристики одного из термисторов приведен на рис. 10.1.

Номинальное сопротивление термистора — это его сопротивление при определенной температуре (обычно 20°С). Термисторы изготовляют с допустимым отклонением от номинального сопротивления ± 20, 10 и 5%. Номинальные сопротивления различных типов термисторов имеют значения от нескольких ом до нескольких сотен килоом.

Коэффициент температурной чувствительности B — это коэффициент в показателе экспоненты температурной характеристики термистора (10.1). Значение этого коэффициента, зависящее от свойств материала термистора, практически постоянно для данного термистора в рабочем диапазоне температур и для различных типов термисторов находится в пределах от 700 до 15 000 K. Коэффициент температурной чувствительности может

быть найден экспериментально путем измерения сопротивлений термистора при температурах T_0 и T по формуле

$$B = \frac{\ln (R/R_0)}{1/T_0 - 1/T}.$$

Температурный коэффициент сопротивления термистора — это величина, определяемая отношением относительного изменения сопротивления термистора к изменению его температуры:

$$\text{ТК}R|_T = \frac{1}{R}\frac{dR}{dT}. \qquad (10.4)$$

Температурный коэффициент сопротивления зависит от температуры, поэтому его необходимо записывать с индексом, указывающим температуру, при которой имеет место данное значение. Зависимость температурного коэффициента сопротивления от температуры можно получить, использовав уравнения (10.4) и (10.1):

$$\text{ТК}R = -B/T^2.$$

Значения температурного коэффициента сопротивления при комнатной температуре различных термисторов находятся в пределах $-(0{,}8...6{,}0)\ 10^{-2}\ \text{K}^{-1}$.

Коэффициент рассеяния термистора H численно равен мощности, рассеиваемой термистором при разности температур термистора и окружающей среды в 1 К, или, другими словами, численно равен мощности, которую надо выделить в термисторе, чтобы нагреть его на 1 К.

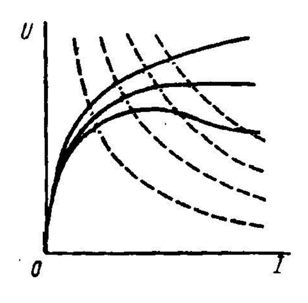


Рис. 10.2. Три типа статических ВАХ термисторов прямого подогрева (сплошные линии) и гиперболы равной мощности (штриховые линии)

Статическая вольт-амперная характеристика термистора — это зависимость падения напряжения на термисторе от проходящего через него тока в условиях теплового равновесия между термистором и окружающей средой.

На рис. 10.2 показаны статические ВАХ термисторов. Линейность характеристик при малых токах и напряжениях объясняется тем, что выделяемая в термисторе мощность недостаточна для существенного изменения его температуры. При увеличении тока, проходящего через термистор, выделяемая в нем мощность повышает его температуру. Таким образом, сопротивление термистора определяется суммарной температурой — температурой окружающей среды и температурой перегрева термистора. При этих токах сопротивление термистора уменьшается с увеличением тока и температуры в соответствии с (10.1), линейность статической ВАХ нарушается. При дальнейшем увеличении тока и большой температурной чувствительности термистора может наблю-

даться падающий участок статической ВАХ, т. е. уменьшение напряжения на термисторе с увеличением проходящего через него тока.

Мощность, выделяющаяся в термисторе, непрерывно возрастает с увеличением проходящего через термистор тока, несмотря на уменьшение падающего на термисторе напряжения. В результате гиперболы равной мощности пересекают статическую ВАХ термистора только в одной точке (рис. 10.2).

Для каждой точки статической ВАХ термистора справедливо уравненне теплового баланса между мощностью, выделяющейся в термисторе из-за проходящего тока, и мощностью, которую он рассеивает в окружающую среду:

$$U^2/R = I^2 R = H(T - T_{\text{окр}}), \tag{10.5}$$

где T — температура термистора; $T_{\text{окр}}$ — температура окружающей среды.

Из уравнения (10.5) с учетом (10.1) можно получить уравнення статической ВАХ термистора в параметрическом виде:

$$\begin{cases} U = \sqrt{HR_\infty (T - T_{\text{окр}}) \exp \frac{B}{T}}; & (10.6) \\ I = \sqrt{\frac{H}{R_\infty} (T - T_{\text{окр}}) \exp \left(-\frac{B}{T}\right)}. & (10.7) \end{cases}$$

Вид статической ВАХ термистора определяется коэффициентом рассеяния H, коэффициентом температурной чувствительности B, номиальным сопротивлеиием термистора и температурой окружающей среды. Так, при уменьшенни коэффицнента рассеяния H (например, при уменьшении давления воздуха, окружающего термнстор) происходит более интенсивный разогрев термистора и, следовательно, те же температуры достигаются при меньших мощностях, выделяемых в термисторе при прохождении тока, т. е. статическая ВАХ смещается вниз (в область меньших напряжений).

При увеличении температуры окружающей среды уменьшается сопротнвление термистора, снижается максимум статической ВАХ и уменьшается ее крутизна. Такую зависимость используют в системах автоматического контроля и регулироваиия температуры.

Увеличение коэффициеита температурной чувствительности B приводит к смещению максимума статической ВАХ в сторону меньших мощностей, а крутизна падающего участка возрастает.

Исследуем уравнение (10.6) на максимум функции, полагая коэффициент температурной чувствительности и коэффициеит рассеяния постоянными. Для этого первую производиую от напряжения по току приравняем нулю. В связи с параметрическим видом ВАХ в данном случае производную от напряжеиия по вспомогательной переменной, т. е. по параметру T, умножим на

производную температуры по току и приравияем полученное произведение нулю:

$$\frac{dU}{dI}=\frac{dU}{dT}\,\frac{dT}{dI}=0.$$

В результате получим

$$T^2_{U\,\text{экстр}}-BT_{U\,\text{экстр}}+BT_{\text{окр}}=0; \qquad (10.8)$$

отсюда

$$T_{U\,\text{экстр}}=\frac{B\pm\sqrt{B(B-4T_{\text{окр}})}}{2}. \qquad (10.9)$$

Из уравнения (10.9) следует:

1. Статическая ВАХ термистора будет иметь экстремальные значения напряжения (кривая *3* на рис. 10.2) только при условии

$$B>4T_{\text{окр}}. \qquad (10.10)$$

2. Существует два решения, соответствующие максимуму и минимуму напряжения на термисторе. Однако при минимуме напряжения на термисторе температура его оказывается выше допустимой, т. е. практически минимум напряжения находится за пределами рабочего диапазона температур термистора.

3. Температура, а значит, и сопротивление термистора при экстремальных значениях напряжения определяются только значениями B и $T_{\text{окр}}$. Температура термистора при экстремальных значениях напряжения не зависит, например, от коэффициента рассеяния. Поэтому максимумы (и минимумы) статических ВАХ термистора, помещенного в различные среды, должны наблюдаться при одних и тех же сопротивлениях термистора. Коэффициент рассеяния сказывается только на значениях напряжения и тока в точке максимума (и минимума) статической ВАХ, что следует из уравнений (10.6) и (10.9).

Максимально допустимая температура термистора — это температура, при которой еще не происходит необратимых изменений параметров и характеристик термистора. Максимально допустимая температура определяется не только свойствами исходных материалов термистора, но и его конструктивными особенностями.

Максимально допустимая мощность рассеяния термистора — это мощность, при которой термистор, находящийся в спокойном воздухе при температуре 20°С, разогревается при прохождении тока до максимально допустимой температуры. При уменьшении температуры окружающего воздуха, а также при работе термистора в средах, обеспечивающих лучший теплоотвод, мощность рассеяния может превышать максимально допустимое значение.

Коэффициент энергетической чувствительности термистора G численно равен мощности, которую необходимо подвести к термистору для уменьшения его сопротивления на 1%. Коэффициент

энергетической чувствительности связан с коэффициентом рассеяния и температурным коэффициентом сопротивления соотношением

$$G = H/\text{TK}R.$$

Значение коэффициента энергетической чувствительности зависит от режима работы термистора, т. е. оно различно в каждой точке статической ВАХ.

Постоянная времени термистора — это время, в течение которого температура термистора уменьшится на 63% (в e раз) по отношению к разности температур термистора и окружающей среды (например, при переносе термистора из воздушной среды с температурой 120°С в воздушную среду с температурой 20°С). Тепловая инерционность термистора, характеризуемая его постоянной времени, определяется конструкцией и размерами термистора и зависит от теплопроводности среды, в которой находится термистор. Для разных типов термисторов постоянная времени лежит в пределах от 0,5 до 140 с.

§ 10.2. БОЛОМЕТРЫ

Полупроводниковый болометр — это прибор, предназначенный для индикации и измерения теплового излучения (оптического или инфракрасного диапазона частот электромагнитного излучения).

Обычно болометр состоит из двух пленочных термисторов (толщиной до 10 мкм). Один из термисторов болометра является активным, т. е. непосредственно подвергается воздействию измеряемого излучения. Сопротивление этого термистора изменяется в результате нагрева при облучении электромагнитным излучением оптического или инфракрасного диапазона частот. Второй термистор — компенсационный — служит для компенсации возможных изменений температуры окружающей среды. Компенсационный термистор должен быть экранирован от измеряемого излучения. Активный и компенсационный термисторы помещают в один герметичный корпус.

Болометры обычно имеют три внешних вывода: от активного и компенсационного термисторов и от средней точки.

Для характеристики болометров используют следующие параметры: 1) сопротивление активного термистора болометра при комнатной температуре; 2) рабочее напряжение; 3) чувствительность при определенной частоте модуляции лучистого потока, равная отношению полезного сигнала, снимаемого с болометра на вход усилителя, к мощности излучения, падающего на болометр; 4) порог чувствительности, численно равный мощности излучения, которая вызывает сигнал, эквивалентный уровню собственных шумов болометра, т. е. порог чувствительности определяется минимальной мощностью излучения, которую при данных условиях способен зарегистрировать болометр; 5) по-

стоянная времени, характеризующая тепловую инерционность активного термистора; 6) уровень собственных шумов.

Полупроводниковые болометры применяют в различных системах ориентации, для бесконтактного и дистанционного измерения температур и т. д.

§ 10.3. ТЕРМИСТОРЫ КОСВЕННОГО ПОДОГРЕВА

Термистор косвенного подогрева — это термистор, имеющий дополнительный источник теплоты — подогреватель.

Конструктивное исполнение термисторов косвенного подогрева может быть различным. Часто подогреватель делают в виде обмотки на изоляционной трубке, внутри которой расположен термистор. В других случаях сам термистор сделан в виде трубки, внутри которой проходит нить подогрева. Общим для всех

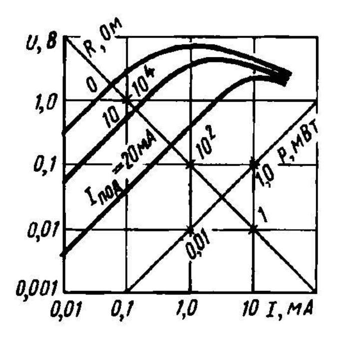


Рис. 10.3. Статические ВАХ термистора косвенного подогрева при различных токах через подогреватель

Рис. 10.4. Подогревная характеристика термистора косвенного подогрева

конструкций термисторов косвенного подогрева является наличие у них двух электрически изолированных друг от друга цепей: управляющей и управляемой.

Кроме таких параметров, как номинальное сопротивление и коэффициент температурной чувствительности, термисторы косвенного подогрева имеют свои специфические характеристики и параметры.

Статические вольт-амперные характеристики термистора косвенного подогрева приводят для различных токов через подогреватель (рис. 10.3).

Подогревная характеристика — это зависимость сопротивления термистора косвенного подогрева от мощности, выделяемой в спирали подогревной обмотки (рис. 10.4).

Для получения наибольшей чувствительности термистора косвенного подогрева, т. е. наибольшего изменения сопротивле-

ния, его следует использовать в режимах, при которых мощностью, выделяемой в самом термочувствительном элементе проходящим через него током, можно было бы пренебречь.

Коэффициент тепловой связи — это отношение мощности $P_т$, необходимой для разогрева термочувствительного элемента до некоторой температуры при прямом нагреве, к мощности $P_{под}$, необходимой для разогрева до той же температуры при косвенном подогреве, т. е. путем пропускания тока через подогреватель: $K = P_т/P_{под}$.

Обычно для определения коэффициента тепловой связи термистор косвенного подогрева разогревают до так называемого горячего сопротивления термистора при максимальной мощности, выделяемой в подогревателе. Коэффициент тепловой связи обычно составляет 0,5...0,97, т. е. меньше единицы, так как часть теплоты, выделяемая подогревателем, неизбежно теряется.

Постоянные времени. Тепловая инерционность термисторов косвенного подогрева характеризуется двумя постоянными времени. За первую постоянную времени принимают время, в течение которого температура термочувствительного элемента изменяется в *e* раз по отношению к установившемуся значению при мгновенном изменении мощности в цепи подогревателя. Вторая постоянная времени характеризует задержку в изменении температуры термочувствительного элемента по отношению к изменению температуры подогревателя. Таким образом, первая постоянная времени характеризует тепловую инерционность всей конструкции термистора косвенного подогрева; вторая постоянная времени — тепловую инерционность термочувствительного элемента.

§ 10.4. ПОЗИСТОРЫ

Позистор — это полупроводниковый терморезистор с положительным температурным коэффициентом сопротивления.

В массовом производстве позисторы делают на основе керамики из титаната бария. Титанат бария $BaTiO_3$ — диэлектрик, известный с начала 40-х годов нашего столетия, с удельным сопротивлением при комнатной температуре $10^{10}...10^{12}$ Ом·см, что значительно превышает удельное сопротивление полупроводников. Если же в состав керамики из титаната бария ввести примеси редкоземельных элементов (лантана, церия или др.) либо других элементов (ниобия, тантала, сурьмы, висмута и т. п.), имеющих валентность большую, чем у титана, и ионный радиус, близкий к радиусу иона титана, то это приведет к уменьшению удельного сопротивления до $10...10^2$ Ом·см, что соответствует удельному сопротивлению полупроводниковых материалов (рис. 10.5). Полупроводниковый титанат бария обладает аномальной температурной зависимостью удельного сопротивления: в узком диапазоне температур при нагреве выше

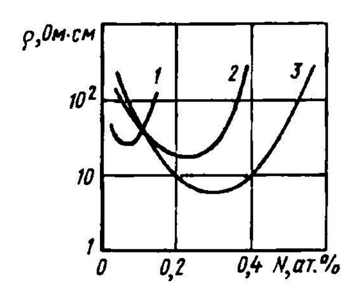


Рис. 10.5. Зависимости удельного сопротивления титаната бария от концентрации различных примесей:

1 — неодим; 2 — церий, лантан, ниобий; 3 — иттрий

точки Кюри удельное сопротивление полупроводникового титаната бария увеличивается на несколько порядков.

Механизм электропроводности полупроводникового титаната бария при наличии примесей можно представить следующим образом. Примесь редкоземельного элемента (например, лантана) замещает в узле кристаллической решетки барий. Часть атомов титана, поддерживая электрическую нейтральность всего кристалла, захватывает лишние валентные электроны лантана, имеющего большую валентность, чем валентность бария. Захватываемые электроны, находясь в квазиустойчивом состоянии, легко перемещаются под действием электрического поля и обусловливают электропроводность материала. Сказанное можно проиллюстрировать следующей формулой:

$$Ba^{2+}Ti^{4+}O_3^{2-} + xLa^{3+} \rightarrow (Ba^{2+}_{1-x}La^{3+}_x)(Ti^{4+}_{1-x}Ti^{3+}_x)O_3^{2-}$$

При замещении титана в титанате бария примесями другого элемента (например, атомами тантала) аналогичный процесс электропроводности можно представить так:

$$Ba^{2+}Ti^{4+}O_3^{2-} + xTa^{5+} \rightarrow Ba^{2+}(Ti^{4+}_{1-2x}Ti^{3+}_xTa^{5+}_x)O_3^{2-}$$

Полупроводниковые свойства керамики на основе титаната бария можно получить также методом частичного восстановления:

$$Ba^{2+}Ti^{4+}O_3^{2-} - xO^{2-} \rightarrow Ba^{2+}(Ti^{4+}_{1-2x}Ti^{3+}_{2x})O^{2-}_{3-x}$$

Во всех перечисленных случаях в полупроводниковом титанате бария существуют четырехвалентные и трехвалентные ионы титана. Между разновалентиыми ионами титана может происходить обмен электронами. При этом каждый ион титана становится то трех-, то четырехвалентным (рис. 10.6). Этот процесс является причиной электропроводности полупроводникового титаната бария.

Появление полупроводниковых свойств в ионных кристаллах под влиянием примесей наблюдается также и для оксида никеля. Полупроводники, изготовляемые подобным методом, иногда называют полупроводниками с управляемой валентностью.

Технология изготовления пози-

La^{3+} Ba^{2+} Ba^{2+}

$Ti^{3+(4+)}$ $Ti^{4+(3+)}$

Ba^{2+} Ba^{2+} Ba^{2+}

Рис. 10.6. Схема, поясняющая электропроводность полупроводникового титаната бария путем обмена электронами между соседними ионами титана

сторов аналогична технологии изготовления нзделий из других керамических материалов. После смешивания исходных компонентов и веществ, содержащих примесные элементы, проводят первичный обжиг при температуре около 1000°С. Полученную твердую массу измельчают, а затем формуют заготовки (обычно в виде дисков различных размеров). Вторичный обжиг производят при температуре 1300...1400°С.

Таким образом, резистивный слой позистора состоит из большого числа контактирующих между собой зерен или кристаллитов полупроводникового титаната бария. Сопротивление позистора определяется сопротивленнем обедненных поверхностных слоев на зернах титаната бария. Высота поверхностных потенциальных барьеров оказывается малой при температурах ниже точки Кюри, когда в зернах титаната бария существует спонтанная поляризация и материал обладает очень большой диэлектрической проницаемостью. При температурах, больших точки Кюри, титанат бария претерпевает фазовое превращение из сегнетоэлектрического в параэлектрическое состояние. При этом пропадает спонтанная поляризация, резко уменьшается диэлектрическая проницаемость, растет высота поверхностных потенциальных барьеров на зернах титаната бария и увеличивается сопротивление позистора (рис. 10.7).

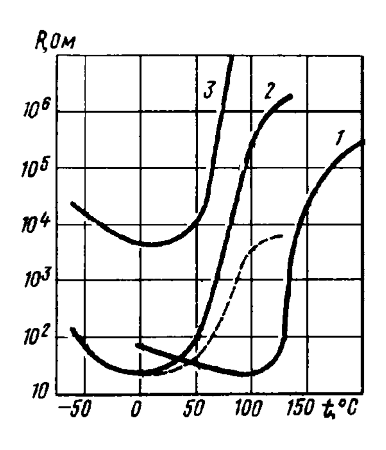


Рис. 10.7. Температурные характеристики некоторых позисторов: 1 — СТ5-1; 2 — СТ6-2Б; 3 — СТ6-3Б; штрнховая кривая для СТ6-2Б получена при нагреве его проходящнм током

Температурные характеристики позисторов связаны с точкой Кюри керамики. Точка Кюри титаната бария может быть смещена в сторону низких температур путем частичного замещения бария стронцием. И наоборот, точка Кюри может быть смещена в сторону больших температур частичной заменой бария свинцом. Уменьшает точку Кюри и частичная замена титана цирконием, оловом или самарием. Такое регулирование позволяет создавать позисторы, у которых положительный температурный коэффициент сопротнвления наблюдается в разных диапазонах температур.

Титанат бария, конечно, не единственный сегнетоэлектрический материал, пригодный для изготовлення позисторов. Положительный температурный коэффициент сопротивления можно получить также при использовании материалов системы $SrTiO_3$ с добавкой ниобия, системы $PbTiO_3$ с добавкой лантана, системы $PbNbO_6$ с добавкой вольфрама и др. Для совершенствования позисторов необходимы исследования материалов с точки

зреиия повторяемости свойств, надежности и увеличения срока службы, т. е. качеств, необходнмых для внедреиия в массовое производство.

Иногда для создания позисторов используют моиокристаллические кремиий, германий и другие полупроводииковые материалы. Прииицип действия таких позисторов основан иа уменьшеиии подвижиости иосителей заряда с увеличеиием температуры в результате увеличения их рассеяиия иа тепловых колебаииях атомов кристаллической решетки. Так, позисторы, изготовлеиные из моиокристаллического кремиия с малой коицеитрацией примесей ($10^{15}...10^{17}$ см $^{-3}$), могут иметь температуриый коэффициеит сопротивлеиия $(0,7...1) \cdot 10^{-2}$ K^{-1} с положительиым знаком в диапазоие температур от 20 до ~ 100°С.

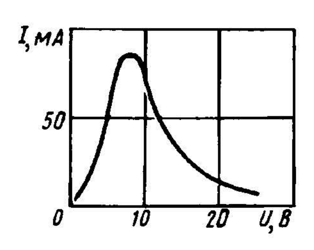


Рис. 10.8. Статнческая ВАХ позистора СТ5-1

Одиим из преимуществ монокристаллических позисторов перед поликристаллическими является возможиость изготовлеиия их с отиосительно малым разбросом параметров и характеристик. При использоваиии для изготовлеиия моиокристаллических позисторов широко распростраиеиной плаиарной технологии можио создавать моиокристаллические позисторы с разбросом номинальных сопротивлений 1...2%. Однако из-за меиьшей стоимости и из-за больших температурных коэффициеитов сопротивлеиия поликристаллические позисторы иашли более широкое примеиеиие.

Свойства позисторов оцениваются характеристиками и параметрами, аиалогичиыми характеристикам и параметрам термисторов с отрицательиым температуриым коэффициеитом сопротивления.

Температурная характеристика. Зависимость сопротивления различных позисторов от температуры показаиа иа рис. 10.7. При отиосительио малых и больших температурах температуриые характеристики позисторов соответствуют температуриым зависимостям удельиого сопротивления обычиых иоииых полупроводииков, т. е. в этих диапазоиах температур позисторы имеют отрицательиый температуриый коэффициеит сопротивления.

Существеиным иедостатком поликристаллических позисторов является иеоднозиачиость температуриых характеристик при различиых иапряжеииях, приложеииых к позистору: увеличеиие иапряжеиия сиижает сопротивлеиие и умеиьшает измеиеиие сопротивлеиия при разогреве позистора (рис. 10.7).

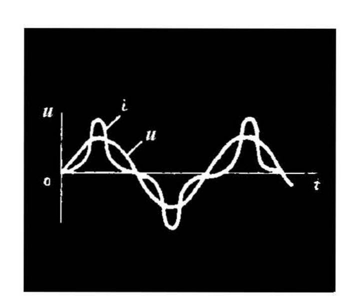
Температурный коэффициент сопротивления для позисторов оказывается ие очеиь удобиым параметром, так как его значение сильио зависит от температуры.

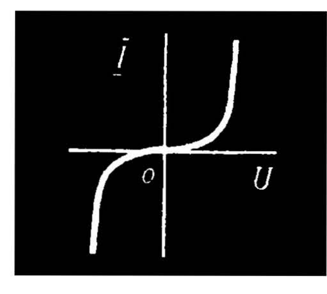
Статическая вольт-амперная характеристика позистора

(рис. 10.8), так же как и ВАХ термистора, представляет собой зависимость напряжения на позисторе от проходящего через него тока при условии теплового равновесия между теплотой, выделяемой в позисторе, и теплотой, отводимой от него.

Контрольные вопросы

1. Какие физические явления могут быть положены в основу создания термисторов?
2. При каких условиях термистор обладает отрицательным дифференциальным сопротивлением?
3. Что характеризует постоянная времени термистора?
4. Почему полупроводниковые болометры обычно состоят из двух термисторов?
5. На каких физических явлениях может быть основан принцип действия позисторов?





глава

11

Варисторы

§ 11.1. ПРИНЦИП ДЕЙСТВИЯ ВАРИСТОРОВ ИЗ КАРБИДА КРЕМНИЯ

Варнстор — это полупроводннковый резнстор, сопротнвленне которого завнснт от прнложенного напряження.

Варисторы изготовляют методом керамической техиологии, т. е. путем высокотемпературного обжига заготовок из порошкообразиого карбида кремния со связкой, в качестве которой обычно используют глииу. Конструктивное оформлеиие варисторов может быть различиым в зависимости от иазначеиия и необходимых параметров. Чаще всего варисторы имеют форму стержней или дисков с электродами иа противоположиых концах или поверхиостях, к которым припаяны выводы для присоедииеиия к схеме.

Нелииейность ВАХ варисторов (рис. 11.1) обусловлена явлеииями на точечных контактах между кристаллами карбида кремиия. При малых напряжеииях иа варисторе может происходить туииелироваиие электроиов сквозь тоикие потенциальиые барьеры, существующие на поверхности кристаллов карбида кремиия.

При больших иапряжениях на варисторе и соответствеиио при больших токах, проходящих через него, плотность тока в точечиых коитактах оказывается очень большой. Все напряжеиие, приложеииое к варистору, падает иа точечиых контактах. Поэтому удельиая мощность (мощиость в едииице объема), выделяющаяся в точечных коитактах, достигает таких значеиий, которые иельзя ие учитывать. Разогрев точечных контактов приводит к умеиьшению их сопротивления и к нелииейности ВАХ.

Сопротивлеиие точечиых коитактов определяется сопротивлением растекаиия (см. § 3.21), т. е. сопротивлением малых актив-

ных областей полупроводника под точечными контактами. Из-за малости активных областей их разогрев практически не приводит к повышению температуры всего варистора. Кроме того, малые объемы активных областей обеспечивают малую инерционность тепловых процессов — разогрева и охлаждения этих областей. Теоретические расчеты показывают, что тепловая постоянная времени активных областей может составлять $10^{-6}...10^{-8}$ с.

Считая разогрев активных областей под точечными контактами одним из основных процессов, приводящих к нелинейности ВАХ в рабочем диапазоне напряжений и токов для варистора, можно получить ряд важных зависимостей и характеристик варисторов.

Температурная зависимость удельной проводимости полупроводников соответствует уравнению

$$\gamma = \gamma_\infty \exp(-B/T). \qquad (11.1).$$

Рис. 11.1. ВАХ одного из варисторов

Сопротивление растекания двух контактирующих кристаллов [см. (3.103)]

$$R_{\text{кон}} \approx \frac{2}{\pi d\gamma} = \frac{2}{\pi d\gamma_\infty}\exp\frac{B}{T}, \qquad (11.2)$$

где d — диаметр точечного контакта; B — коэффициент температурной чувствительности поверхностных слоев кристаллов карбида кремния.

Тогда статическое сопротивление варистора, состоящего из a параллельно включенных цепочек, имеющих, в свою очередь, b последовательно включенных контактирующих кристаллов,

$$R = \frac{b}{a} R_{\text{кон}} = \frac{2b}{a\pi d\gamma_\infty}\exp\frac{B}{T}. \qquad (11.3)$$

Уравнение теплового баланса для активных областей варистора

$$U^2/R = I^2R = H(T - T_0), \qquad (11.4)$$

где H — коэффициент рассеяния активных областей; T — температура активных областей; T_0 — температура среды, окружающей активные области.

Из уравнения (11.4), учитывая (11.3), можно получить уравнения ВАХ варисторов в параметрической форме:

$$\begin{cases} U = \sqrt{\dfrac{2Hb(T-T_0)}{a\pi d\gamma_\infty}\exp\dfrac{B}{T}}; & (11.5) \\ I = \sqrt{\dfrac{a\pi d H\gamma_\infty(T-T_0)}{2b}\exp\left(-\dfrac{B}{T}\right)}. & (11.6) \end{cases}$$

Уравнения (11.5) и (11.6) неудобны для расчета цепей с варисторами, так как содержат ряд величин (a, b, d, H), значения которых практически невозможно определить непосредственно. Однако с помощью этих уравнений можно выяснить некоторые требования, предъявляемые к исходному материалу для изготовления варисторов с заданными свойствами. Кроме того, эти уравнения дают возможность установить связь между различными параметрами варисторов.

§ 11.2. ХАРАКТЕРИСТИКИ

Коэффициент нелинейности варистора — это отношение статического R и дифференциального r сопротивлений при заданном постоянном напряжении на варисторе:

$$\beta = \frac{R}{r} = \frac{U}{I}\frac{dI}{dU}. \qquad (11.7)$$

Из уравнений (11.5) и (11.6) найдем дифференциальное сопротивление варистора:

$$r = \frac{dU}{dI} = \frac{dU}{dT} \Big/ \frac{dT}{dI} = \frac{2b}{ad\pi\gamma_\infty}\left(\frac{T^2 - BT + BT_0}{T^2 + BT - BT_0}\right)\exp\frac{B}{T}. \qquad (11.8)$$

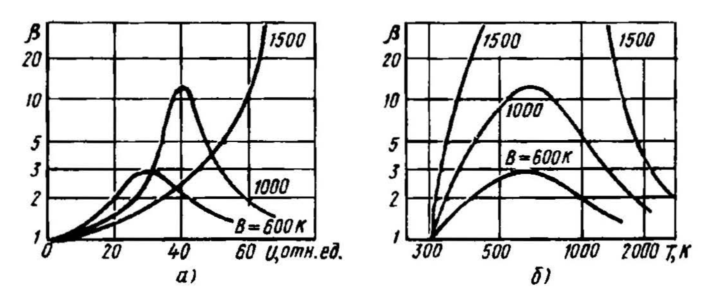


Рис. 11.2. Расчетные зависимости коэффициента нелинейности варисторов из материалов с различным значением коэффициента температурной чувствительности B:

а — от напряжения; *б* — от температуры активной области

Тогда с учетом соотношений (11.7) и (11.8) коэффициент нелинейности варистора

$$\beta = \frac{R}{r} = \frac{T^2 + BT - BT_0}{T^2 - BT + BT_0}. \qquad (11.9)$$

На рис. 11.2 показаны расчетные зависимости коэффициента нелинейности от напряжения и температуры активных областей варисторов из материалов с различными значениями коэффициента температурной чувствительности B поверхностных слоев кристаллов.

Для определения положения максимумов этих зависимостей продифференцируем (11.9) по температуре и приравняем производную нулю. Тогда получим условие $T=2T_0$, при котором коэффициент нелинейности имеет максимальное значение:

$$\beta_{max}=\frac{4T_0+B}{4T_0-B}. \qquad (11.10)$$

Из соотношения (11.10) следует, что при $B<4T_0$ должны наблюдаться нелинейные ВАХ с положительным коэффициентом нелинейности. Максимальный коэффициент нелинейности должен быть равен бесконечности при $B=4T_0$. Если $B>4T_0$, то у варистора должно наблюдаться отрицательное дифференциальное сопротивление, коэффициент нелинейности при этом тоже будет отрицательным.

На основе этих расчетов (рис. 11.2, *б*) можно заключить, что температура активных областей варистора может превышать температуру окружающей среды на несколько сотен градусов. Следовательно, для изготовления варисторов со стабильными параметрами необходим термостойкий материал. Именно поэтому при массовом производстве варисторов используют карбид кремния — один из самых термостойких материалов. Одновременно поликристаллический карбид кремния является очень дешевым материалом.

Основными примесями в техническом карбиде кремния являются азот и алюминий. Энергия ионизации этих примесей в карбиде кремния невелика (особенно при большой концентрации основных и компенсирующих примесей, что имеет место в техническом карбиде кремния), соответственно мало и значение коэффициента температурной чувствительности B. Поэтому коэффициент нелинейности варисторов не превышает 6, что ограничивает возможные применения варисторов.

Увеличение температуры окружающей среды должно привести к уменьшению коэффициента нелинейности (11.9) и незначительному сдвигу максимума кривой $\beta=f(U)$ по напряжению (рис. 11.3).

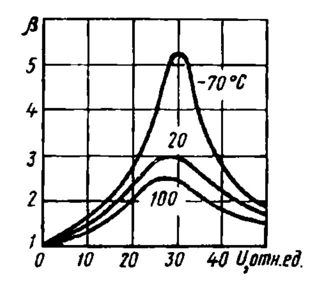


Рис. 11.3. Расчетная зависимость коэффициента нелинейности варистора от напряжения при различных температурах ($B=600$ К)

Вольт-амперная характеристика варистора, как было отмечено, должна соответствовать уравнениям (11.5) и (11.6). Если же варистор работает в узком диапазоне изменений напряжений и токов, то коэффициент нелинейности в этом диапазоне можно считать постоянным:

$$\beta=UdI/(IdU)=\text{const}.$$

Тогда

$$\int dI/I = \beta \int dU/U; \ \ln I = \beta \ln U + \ln A$$

и ВАХ варистора будет соответствовать уравнению

$$I = AU^{\beta}, \tag{11.11}$$

где A — коэффициент, значение которого зависит от типа варистора и от температуры.

Иногда ВАХ варисторов аппроксимируют уравнением

$$U = A_1 I^{\alpha}, \tag{11.12}$$

где $\alpha = 1/\beta$ и $A_1 = A^{-1/\beta} = A^{-\alpha}$.

Используя уравнения (11.11) и (11.12), статическое сопротивление варистора можно выразить как функцию тока или напряжения:

$$R(I) = A_1 I^{1/\beta - 1}; \tag{11.13}$$

$$R(U) = A_1^{\beta} U^{1-\beta}. \tag{11.14}$$

Температурные коэффициенты статического сопротивления, напряжения и тока. В связи с нелинейностью ВАХ следует различать температурные коэффициенты статического сопротивления варистора, измеренные при постоянных напряжении или токе, а также температурные коэффициенты напряжения и тока. Из уравнений (11.11) — (11.14) с учетом температурного изменения коэффициентов A и A_1 получим:

$$\text{ТК}R|_{U=\text{const}} = \frac{1}{R}\frac{\partial R}{\partial T}\bigg|_{U=\text{const}} =$$

$$= \text{ТК}A_1 + \left(\frac{1}{\beta} - 1\right)\text{ТК}I|_{U=\text{const}}; \tag{11.15}$$

$$\text{ТК}R|_{I=\text{const}} = \frac{1}{R}\frac{\partial R}{\partial T}\bigg|_{I=\text{const}} =$$

$$= \beta\text{ТК}A_1 + (1-\beta)\text{ТК}U|_{I=\text{const}}; \tag{11.16}$$

$$\text{ТК}I|_{U=\text{const}} = \frac{1}{I}\frac{\partial I}{\partial T}\bigg|_{U=\text{const}} = -\beta\text{ТК}A_1; \tag{11.17}$$

$$\text{ТК}U|_{I=\text{const}} = \frac{1}{U}\frac{\partial U}{\partial T}\bigg|_{I=\text{const}} = \text{ТК}A_1. \tag{11.18}$$

При малых напряжениях на варисторе, когда коэффициент нелинейности $\beta = 1$, т. е. на линейном участке ВАХ

$$\text{ТК}A_1 = \frac{1}{A_1}\frac{dA_1}{dT} = \text{ТК}R|_{U\approx 0} = -\frac{B}{T^2}. \tag{11.19}$$

Используя уравнения (11.15) — (11.18), определим соотношения между различными температурными коэффициентами варистора:

$$\text{TK}R|_{U=\text{const}} = \beta \text{TK}R|_{I=\text{const}} ;$$

$$\text{TK}I|_{U=\text{const}} = -\beta \text{TK}U|_{I=\text{const}} .$$

У варисторов, выпускаемых отечественной промышленностью, в диапазоне температур от —40 до +100°C

$$-\text{TK}R|_{U=\text{const}} = \text{TK}I|_{U=\text{const}} < 7 \cdot 10^{-3}\text{K}^{-1} ;$$

$$-\text{TK}R|_{I=\text{const}} = -\text{TK}U|_{I=\text{const}} < 1{,}4 \cdot 10^{-3}\text{K}^{-1} .$$

Из уравнений (11.15)...(11.19) и (11.10) следует, что у варисторов с бо́льшим значением максимального коэффициента нелинейности должны быть бо́льшие по абсолютному значению температурные коэффициенты сопротивления, тока и напряжения. Между максимальным коэффициентом нелинейности и температурными коэффициентами сопротивления, тока и напряжения должна быть взаимосвязь, так как они зависят от коэффициента температурной чувствительности B:

$$\text{TK}I|_{U=\text{const}} = \frac{4\beta_{\max}(\beta_{\max}-1)}{(\beta_{\max}+1)T_0} . \qquad (11.20)$$

Экспериментально такая взаимосвязь наблюдается, но расчет температурного коэффициента тока варистора по уравнению (11.20) дает несколько завышенные результаты. Числовые расхождения можно объяснить, во-первых, наличием добавочных сопротивлений, включенных последовательно и параллельно сопротивлениям активных областей варистора, и, во-вторых, тем, что температура среды, окружающей активные области варистора, несколько выше температуры среды, окружающей весь варистор.

Частотные свойства варисторов могут определяться либо инерционностью процессов, приводящих к нелинейности ВАХ, либо собственной емкостью варистора. Инерционность разогрева и охлаждения активных областей под точечными контактами между кристаллами очень мала. Поэтому частотные свойства варисторов определяются временем перезаряда их собственной емкости.

§ 11.3. ВАРИСТОРЫ ИЗ ОКСИДНЫХ ПОЛУПРОВОДНИКОВ

Исходным материалом для таких варисторов является оксид цинка ZnO, легированный примесями висмута, кобальта и других элементов. Варисторы из оксидных полупроводников делают методом керамической технологии, которая, однако, имеет ряд особенностей, отличающих ее от технологии производства электроизоляционных изделий. Особенности эти обусловлены тем, что

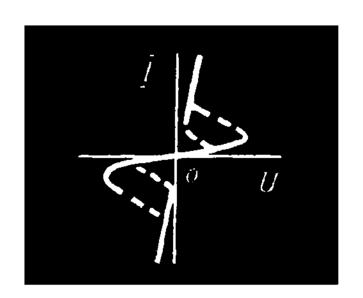
нелинейность ВАХ варисторов из оксидных полупроводников связана не со свойствами основной составляющей полупроводниковой керамики — кристаллитами, а со свойствами межкристаллитных прослоек и потенциальных барьеров на поверхности кристаллитов. Поэтому кроме традиционных требований обеспечения достаточной плотности с минимальной пористостью обожженного материала при изготовлении варисторов надо обеспечить высокоомность межкристаллитного слоя, сопротивление которого должно превышать сопротивление объема кристаллитов.

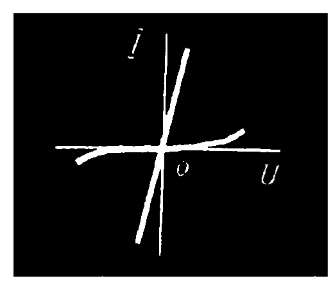
Несмотря на то что первые исследования нелинейных ВАХ кристаллов оксида цинка были проведены еще в 20-х годах О. В. Лосевым в Нижегородской радиолаборатории, а технология первых варисторов из оксидных полупроводников была разработана в 60-х годах, до сих пор нет четкого понимания принципа действия этих варисторов. Поэтому технология их изготовления имеет эмпирический характер.

Отличием варисторов из оксидных полупроводников от варисторов из карбида кремния является возможность относительно просто получить большую нелинейность ВАХ. Коэффициент нелинейности варисторов из оксида цинка может составлять несколько десятков. Отрицательной особенностью варисторов из оксидных полупроводников является их меньшая стабильность как при работе, так и при хранении. Именно поэтому варисторы из оксидных полупроводников до сих пор не получили широкого распространения.

Контрольные вопросы

1. Какой физический процесс можно считать основным при работе варисторов из карбида кремния в диапазоне средних и больших токов?
2. Какими уравнениями можно аппроксимировать ВАХ варистора из карбида кремния? Когда удобнее использовать ту или иную форму аппроксимации?
3. Что такое коэффициент нелинейности варистора?
4. Почему у варисторов из карбида кремния с большим коэффициентом нелинейности наблюдается больший температурный коэффициент сопротивления?
5. Почему у одного и того же варистора различны значения температурных коэффициентов сопротивления, измеренные при постоянных напряжении или токе?





глава

12

Полупроводниковые приборы на аморфных полупроводниках

§ 12.1. ПЕРЕКЛЮЧАТЕЛИ НА АМОРФНЫХ ПОЛУПРОВОДНИКАХ

Переключатель на аморфном полупроводнике — это полупроводниковый прибор, который имеет симметричную вольт-амперную характеристику с участками отрицательного дифференциального сопротивления, т. е. может находиться в закрытом (высокоомном) и открытом (низкоомном) состояниях, но для поддержания открытого состояния через прибор должен непрерывно проходить достаточно большой ток.

Структура переключателей

Аморфные, или стеклообразные, полупроводники можно разделить на *оксидные, элементные и халькогенидные* (или бескислородные). Оксидные аморфные полупроводники или стекла состоят из оксидов различных элементов. Элементные аморфные полупроводники или элементные стекла могут быть изготовлены из таких элементов, как селен, сера, теллур, германий и фосфор. Из элементных стекол наиболее широко применяется аморфный селен.

Халькогенидные стекла состоят из сульфидов, селенидов и теллуридов различных элементов, т. е. из соединений, аналогичных оксидам, в которых роль кислорода выполняют его аналоги — сера, селен и теллур.

Переключатели делают обычно на основе халькогенидных стекол, например, такого состава $Te_{48}As_{30}Ge_{10}Si_{12}$. Халькогенидное стекло наносят на полированную подложку из графита методом вакуумного осаждения при испарении. Толщина пленки халькогенидного стекла от единиц до сотен микрометров. Одним электродом переключателя является графитовая подложка. Вто-

рым электродом служит тонкий слой металла, нанесенного на пленку халькогенидного стекла также методом испарения в вакууме с последующей конденсацией.

Механизм переключения

Переключатель может находиться в двух состояниях: в закрытом состоянии с большим сопротивлением и в открытом состоянии с малым сопротивлением (рис. 12.1). Переключение в открытое состояние происходит при некотором напряжении переключения $U_п$. При переключении прибор обладает отрицательным дифференциальным сопротивлением (ОДС). В открытом состоянии переключатель находится до тех пор, пока через него идет ток, превышающий ток выключения $I_{выкл}$.

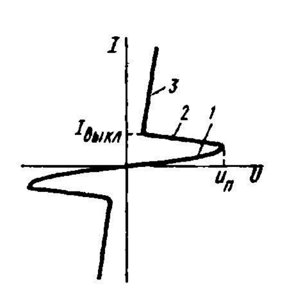


Рис. 12.1. ВАХ переключателя на аморфном полупроводнике:

1 — участок, соответствующий закрытому состоянию переключателя; 2 — переходный участок; 3 — участок, соответствующий открытому состоянию переключателя

Механизм переключения связан с разогревом пленки аморфного полупроводника под действием выделяющейся мощности. Явления, происходящие при переключении, аналогичны явлениям теплового пробоя (см. § 3.13) и явлениям в термисторах с отрицательным температурным коэффициентом сопротивления (см. § 10.1). Следовательно, отрицательное дифференциальное сопротивление и эффект переключения в пленках аморфных полупроводников будут существовать, если удельная проводимость аморфного полупроводника растет с увеличением температуры и если коэффициент температурной чувствительности $B > 4T_0$ (см. § 10.1).

В связи с аналогией в принципах действия переключателей на аморфных полупроводниках и термисторов ВАХ переключателя может быть аппроксимирована той же системой уравнений (10.6) и (10.7).

Разница в механизмах действия переключателей и термисторов связана со значительно меньшим объемом нагреваемой области в аморфной пленке переключателя — проводящего канала или шнура, образующегося при шнуровании тока (см. § 3.13). Проводящий канал возникает перед переключением прибора в открытое состояние и из-за разогрева имеет большую удельную проводимость, чем остальная пассивная часть пленки аморфного полупроводника.

Поэтому величины, входящие в уравнения (10.6) и (10.7), являются параметрами проводящего канала в пленке аморфного полупроводника, т. е. H — это коэффициент рассеяния проводящего канала; T — температура проводящего канала; R_∞ — коэффициент, зависящий от площади сечения проводящего канала,

толщины пленки аморфного полупроводника и его свойств; B — коэффициент температурной чувствительности аморфного полупроводника.

В открытом состоянии переключателя, т. е. при малом сопротивлении проводящего канала, необходимо учитывать еще и сопротивление растекания в графитовой подложке под проводящим каналом. Так, при удельном сопротивлении графита $\varrho = 0{,}2$ Ом·см и радиусе сечения проводящего канала $r = 1$ мкм сопротивление растекания R_s в графитовой подложке, вычисленное по формуле (3.103), составляет 320 Ом. Такое сопротивление может существенно повлиять на ВАХ переключателя в открытом состоянии. С учетом падения напряжения на сопротивлении растекания ВАХ переключателя должна соответствовать уравнению

$$U = \sqrt{HR_{\infty}(T - T_0)\exp\frac{B}{T}} + R_s I \; .$$

Результаты числового расчета по этому уравнению показывают, что значения тока и температуры проводящего канала при минимальном напряжении на переключателе в открытом состоянии в сильной степени зависят от значения сопротивления растекания в графитовой подложке. Например, без учета сопротивления растекания расчетные значения тока и температуры при минимальном напряжении оказываются равными соответственно 1165 мА и 4400 К, что является нереальным и не соответствует экспериментальным значениям. С учетом сопротивления растекания ($R_s = 250$ Ом) те же величины равны 2,2 мА и 430 К, что вполне допустимо. При расчетах в этом примере принято $B = 5250$ К, $HR_{\infty} = 3{,}7 \cdot 10^{-6}$ Ом·Вт·К$^{-1}$.

Удельная проводимость разогретого проводящего канала превышает удельную проводимость остальной пассивной части пленки аморфного полупроводника. Однако площадь поперечного сечения проводящего канала на несколько порядков меньше площади пассивной части пленки аморфного полупроводника между электродами. Поэтому полная проводимость пассивной части пленки может быть больше абсолютного значения дифференциальной проводимости, характерной для переходного участка ВАХ проводящего канала. В таком случае переключатель будет иметь так называемую у-образную ВАХ (см. § 3.13).

Таким образом, переключатели на аморфных полупроводниках могут иметь у-образные ВАХ, которые на первый взгляд противоречат тепловому механизму переключения.

Переходный участок у-образной ВАХ нельзя экспериментально исследовать по точкам даже при использовании идеального генератора тока в качестве источника питания, т. е. источника питания с бесконечно большим собственным сопротивлением.

Напряжение переключения — это минимальное напряжение, при котором происходит переключение переключателя из закрытого в открытое состояние. Для различных переключателей на аморфных полупроводниках напряжение переключения составляет от единиц до десятков вольт.

Напряжение переключения переключателей на аморфных полупроводниках уменьшается с увеличением температуры окружающей среды, так же как и пробивное напряжение при тепловом пробое (см. § 3.13). Однако в переключателях с малой толщиной пленки аморфного полупроводника (несколько микрометров) из-за большой напряженности электрического поля тепловому пробою может предшествовать лавинный пробой. Пробивное напряжение при лавинном пробое растет с увеличением температуры окружающей среды (см. § 3.11). Поэтому у переключателей с малой толщиной пленки аморфного полупроводника может быть сложная зависимость напряжения переключения от температуры. Тем не менее механизм переключения из закрытого состояния в открытое связан только с тепловым пробоем.

Ток выключения — это минимальный ток, при котором переключатель еще находится в открытом состоянии.

При работе переключателя на переменном напряжении или в импульсном режиме необходимо учитывать инерционность процесса разогрева и охлаждения проводящего канала. Ток через переключатель, зависящий от температуры проводящего канала, будет запаздывать относительно приложенного напряжения, т. е. будет сдвинут по фазе. Однако тепловые постоянные времени в связи с малостью объема проводящего канала оказываются также очень малыми ($10^{-6}...10^{-8}$ с). Следовательно, переключатели на аморфных полупроводниках могут работать при частотах до десятков, а иногда и сотен мегагерц.

На время переключения влияют собственная емкость переключателя и емкость внешних элементов схемы. При переключении прибора из закрытого в открытое состояние разряд собственной емкости и емкостей, подключенных параллельно переключателю, происходит через проводящий канал, ускоряя его разогрев и уменьшая тем самым время переключения.

§ 12.2. ЭЛЕМЕНТЫ ПАМЯТИ НА АМОРФНЫХ ПОЛУПРОВОДНИКАХ

Элемент памяти на аморфном полупроводнике — это полупроводниковый прибор-переключатель, сохраняющий открытое состояние после выключения тока и переходящий в закрытое состояние только после прохождения короткого, но относительно мощного импульса тока.

Элементы памяти на аморфных полупроводниках имеют ту же конструкцию, что и переключатели, но в качестве аморфного полупроводника используют обычно халькогенидные стекла трой-

ной системы $Te_{81}Ge_{15}X_4$, где X — ионы мышьяка, серы или сурьмы. Эти соединения при относительно медленном охлаждении в диапазоне температур размягчения обладают склонностью к кристаллизации, а при быстром охлаждении до температур ниже температур размягчения они оказываются в аморфном состоянии. Удельная проводимость полупроводника в поликристаллическом состоянии на несколько порядков выше удельной проводимости того же полупроводника в аморфном состоянии. Поэтому элемент памяти обладает двумя ВАХ, соответствующими открытому (низкоомному) и закрытому (высокоомному) состояниям (рис. 12.2). И то, и другое состояния элемента памяти сохраняются после отключения его от источника питания, т. е. такой прибор является энергонезависимым элементом памяти.

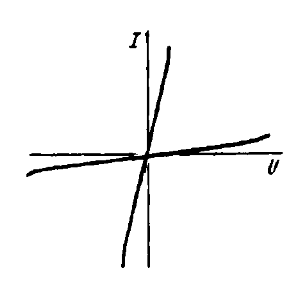


Рис. 12.2. ВАХ запоминающего устройства на аморфном полупроводнике (элемента памяти)

Механизм переключения элемента памяти из закрытого в открытое состояние, так же как и в переключателях на аморфных полупроводниках, связан с разогревом шнура или проводящего канала при тепловом пробое полупроводника, т. е. при напряжении переключения. Для «запоминания» открытого состояния необходимо, чтобы за время охлаждения после выключения тока материал шнура частично или весь успел закристаллизоваться. Кристаллизация произойдет, если, во-первых, температура в шнуре превышает температуру кристаллизации. Во-вторых, должен быть разогрет не только шнур, но и прилегающие области. Только тогда время охлаждения шнура будет достаточным для кристаллизации. Для выполнения этих условий импульс тока, переводящий элемент памяти из закрытого в открытое состояние, должен иметь амплитуду, превышающую 0,1...1 мА, и длительность более 10^{-3} с.

Для перевода элемента памяти из открытого в закрытое состояние необходимо другим импульсом тока разогреть закристаллизованный шнур до температуры плавления, а затем быстро охладить до температуры стеклования. Такие условия будут выполнены, если через элемент памяти, находящийся в открытом состоянии, пропустить импульс тока очень малой длительности ($10^{-5}...10^{-7}$ с). За время действия короткого импульса будет разогрет только шнур. Его охлаждение после действия импульса произойдет быстро и материал шнура не успеет закристаллизоваться. Но для достижения температуры плавления амплитуда импульса тока при этом для большинства элементов памяти должна быть не менее 10 мА.

Информацию (логический нуль или логическую единицу), записанную в элементе памяти, можно считывать с помощью маломощных импульсов. Эти импульсы не должны приводить к

увеличению напряжения на элементе памяти до напряжения переключения, если он находится в закрытом состоянии, и не должны создавать через элемент памяти токи, способные перевести его из открытого в закрытое состояние.

§ 12.3. НАДЕЖНОСТЬ, СТАБИЛЬНОСТЬ И СРОК СЛУЖБЫ ПРИБОРОВ НА АМОРФНЫХ ПОЛУПРОВОДНИКАХ

Надежность, стабильность и срок службы приборов на аморфных полупроводниках довольно низки по сравнению с другими полупроводниковыми приборами. Особенно низки эти параметры при прохождении через приборы на аморфных полупроводниках тока только в одном направлении.

Нестабильность обычно проявляется в уменьшении со временем напряжения переключения и с переходом прибора на аморфных полупроводниках в состояние памяти, т. е. в открытое состояние с эффектом памяти. Связано это явление с постепенной кристаллизацией аморфного полупроводника в проводящем канале, которая происходит тем быстрее, чем выше температура и напряженность поля в проводящем канале. При этих условиях облегчаются дрейф ионов аморфного полупроводника в проводящем канале и их перегруппировка, способствующая кристаллизации.

Число переключений, которое может выдержать переключатель, зависит от тока, проходящего через открытый переключатель, и от длительности прохождения этого тока. Чем больше амплитуда и длительность импульсов тока, тем за меньшее число переключений происходит перегруппировка ионов в проводящем канале, приводящая к кристаллизации в нем аморфного полупроводника.

Максимальное число переключений, которое выдерживает переключатель до выхода его из строя, составляет $10^9...10^{12}$. Для правильной оценки такого, казалось бы большего, числа срабатываний переключателя необходимо учесть, что время рассасывания современных транзисторов может составлять всего несколько наносекунд. Это означает, что транзистор, работающий в схеме переключателя, может по крайней мере за несколько часов выполнить более 10^{12} переключений.

Таким образом, малый срок службы приборов на аморфных полупроводниках, связанный с относительно большой ионной составляющей проводимости этих материалов при повышенных температурах, ставит под сомнение их перспективность и значительно ограничивает их области применения.

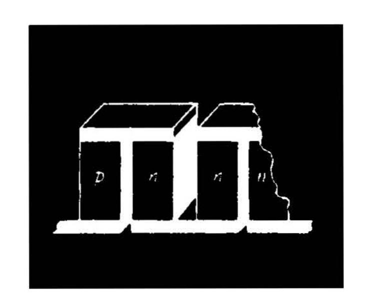
Для повышения стабильности параметров переключателей, их надежности и числа срабатываний до выхода из строя необходимо использовать полупроводниковые материалы с большим коэффициентом температурной чувствительности ($B > 4T_0$) и одно-

временно с более высокой температурой плавления или размягчения, с меньшей ионной составляющей проводимости.

Положительными особенностями приборов на аморфных полупроводниках являются простота их изготовления и высокая устойчивость к проникающей радиации.

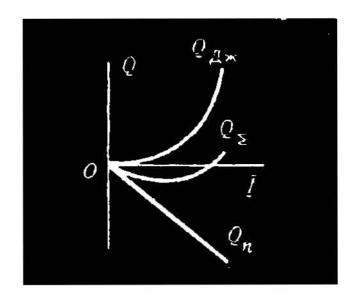
Контрольные вопросы

1. Почему и при каких условиях переключатели на аморфных полупроводниках могут находиться в двух состояниях: закрытом и открытом?
2. Почему у некоторых переключателей не удается исследовать переходный участок ВАХ, соответствующий отрицательному дифференциальному сопротивлению?
3. Как может изменяться при изменении температуры напряжение переключения переключателей на аморфных полупроводниках?
4. Почему в элементах памяти на аморфных полупроводниках открытое состояние может сохраняться даже после отключения прибора от источника питания?



глава

13



Полупроводниковые термоэлектрические устройства

Полупроводниковое термоэлектрическое устройство — это устройство, действие которого основано на использовании термоэлектрических эффектов Пельтье или Зеебека в полупроводниках, предназначенное для получения теплоты или холода с использованием электрической энергии или получения электрической энергии с использованием тепловой.

§ 13.1. КОНСТРУКЦИЯ И ПРИНЦИП ДЕЙСТВИЯ

В полупроводниковых термоэлектрических устройствах используют полупроводниковые *термоэлементы*, каждый из которых состоит из двух *ветвей*, обладающих разным типом электропроводности. Ветвь термоэлемента, материал которого имеет электропроводность *p*-типа, называют *положительной ветвью*. Ветвь термоэлемента с электропроводностью *n*-типа — *отрицательной ветвью*. Положительную и отрицательную ветви полупроводникового термоэлемента соединяют между собой последовательно *контактной пластиной* (рис. 13.1). Зону электрического соединения ветвей полупроводникового термоэлемента называют *спаем*. При работе термоэлемента его спаи имеют различную температуру: один из них является *теплопоглощающим*, другой — *тепловыделяющим*.

Полупроводниковое термоэлектрическое устройство обычно имеет большое число последовательно соединенных между собой термоэлементов в едином конструктивном исполнении — *полупроводниковую термобатарею*. Полупроводниковое термоэлектрическое устройство, включающее одну или несколько электрически

соединенных полупроводниковых термобатарей в едином конструктивном исполнении с системой теплообмена, называют *полупроводниковым термоэлектрическим блоком.*

Возникновение термо-ЭДС (эффект Зеебека)

При наличии разности температур спаев в цепи с термоэлементом появляется *термоэлектродвижущая сила* (термо-ЭДС), которая состоит из трех составляющих.

Первая составляющая термо-ЭДС обусловлена диффузией носителей заряда от нагретого спая, температура которого из-за подводимой к нему тепловой мощности от какого-нибудь источника выше температуры тепловыделяющего спая. Диффузия носителей заряда в ветвях термоэлемента может возникать по двум причинам. Во-первых, в ветвях термоэлемента у нагретого спая оказывается большее число ионизированных примесей. При дополнительной ионизации примесей увеличивается концентрация основных носителей заряда на нагретых концах ветвей термоэлемента. В этом случае возникает диффузия основных носителей в каждой ветви термоэлемента из-за градиента концентрации (рис. 13.2). Во-вторых, если в ветвях термоэлемента все примеси были ионизированы уже при низкой температуре (эта температура выше температуры истощения примесей), то концентрация основных носителей заряда при нагреве практически не увеличится. Но на нагретых концах ветвей термоэлемента носители заряда приобретают бо́льшие энергии. Поэтому опять происходит диффузия основных носителей заряда от нагретого конца в каждой ветви термоэлемента, связанная с выравниванием средней энергии, приходящейся на один носитель определенного знака.

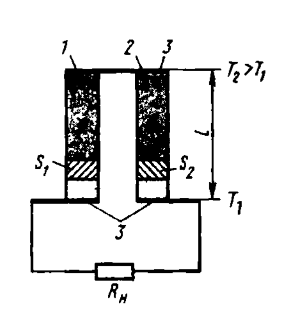


Рис. 13.1. Схема полупроводникового термоэлемента с сопротивлением нагрузки:

1 — положительная ветвь; *2* — отрицательная ветвь; *3* — металлические контактные пластины

Диффузия, например, электронов в отрицательной ветви может происходить только от нагретого конца этой ветви вдоль нее и не может происходить в положительную ветвь, так как переходу электронов в положительную ветвь препятствует потенциальный барьер нагретого спая термоэлемента. Аналогично, дырки в положительной ветви могут диффундировать вдоль нее от нагретого конца.

Перемещение носителей заряда, связанное с их диффузией, нарушает электрическую нейтральность в ветвях термоэлемента — на нагретых концах ветвей термоэлемента остаются нескомпенсированные ионизированные примесные атомы, а на противо-

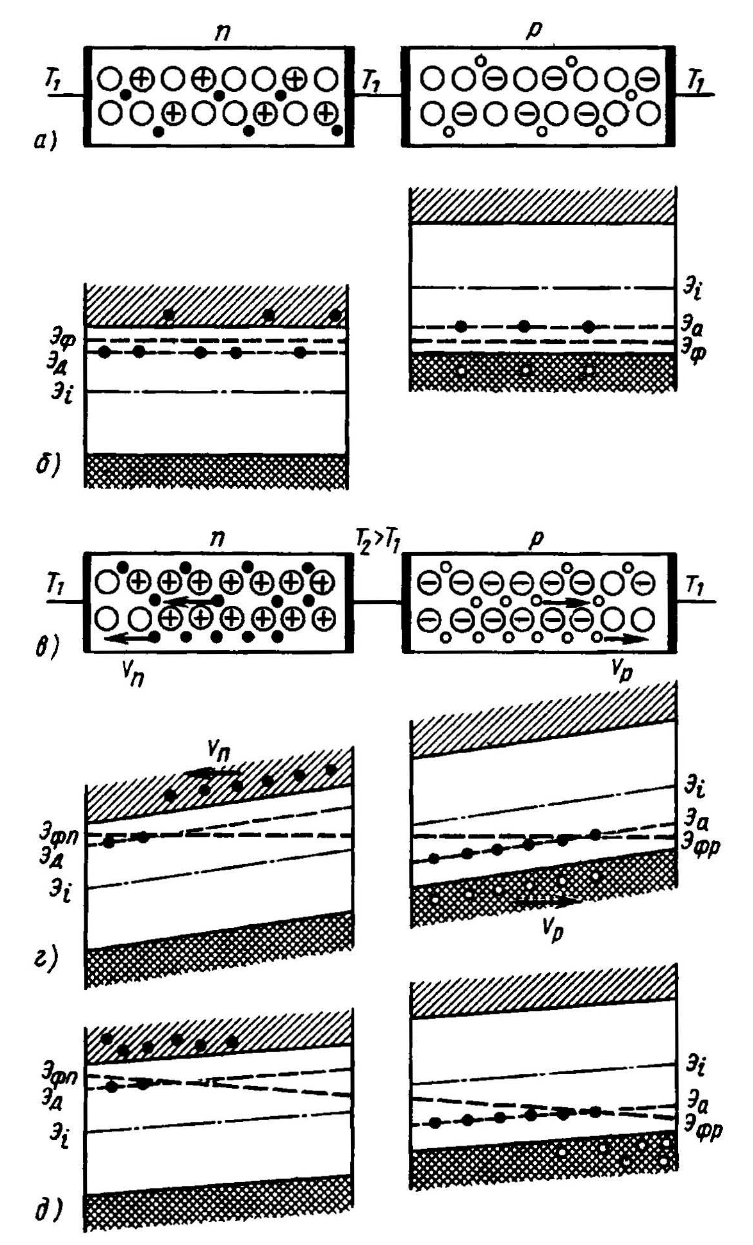


Рис. 13.2. Появление термо-ЭДС в результате диффузии носителей заряда по ветвям термоэлемента:

а — спаи термоэлемента находятся при одной температуре T_1 (не все доноры и акцепторы ионизированы); *б* — энергетические диаграммы ветвей термоэлемента в состоянии термодинамического равновесия (при T_1); *в* — спаи термоэлемента находятся при разных температурах (около нагретых спаев все примеси ионизированы); *г* — энергетические диаграммы ветвей термоэлемента при сопротивлении нагрузки $R_н = 0$ (идет ток, равный току короткого замыкания термоэлемента); *д* — энергетические диаграммы при сопротивлении нагрузки $R_н = \infty$ (возникла термо-ЭДС)

положных концах образуется избыток основных носителей заряда. В результате возникает первая составляющая термо-ЭДС, которую можно назвать диффузионной, так как она возникает из-за процессов диффузии.

Вторая составляющая термо-ЭДС — это следствие температурной зависимости контактной разности потенциалов. Если оба спая термоэлемента имеют одну и ту же температуру, то контакт-

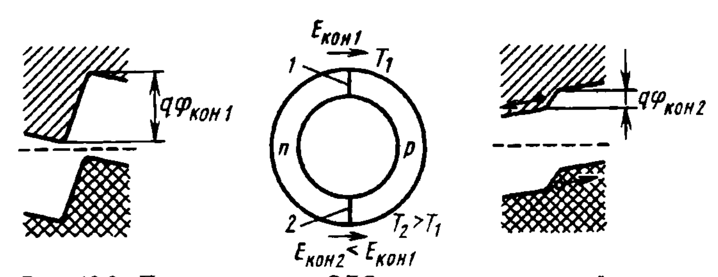


Рис. 13.3. Появление термо-ЭДС из-за температурной зависимости контактной разности потенциалов при коротком замыкании ветвей термоэлемента и энергетические диаграммы спаев, находящихся при разных температурах

ные разности потенциалов на этих спаях равны, направлены в противоположные стороны при обходе контура с термоэлементом и не дают результирующей ЭДС. Если же температура спаев термоэлемента различна, то значение контактиой разности потенциалов на спаях будет также различно (рис. 13.3). Тогда в цепи термоэлемеита появляется вторая составляющая термо-ЭДС с той же полярностью, что и первая составляющая.

Третья составляющая термо-ЭДС возникает в термоэлементе вследствие увлечения носителей заряда квантами тепловой энергии — фононами. Если в ветвях термоэлемента есть градиент температуры, то будет существовать направленное движение фононов от нагретых концов ветвей. В результате столкновений фононов с носителями заряда фононы увлекают за собой электроны в отрицательной ветви и дырки в положительной ветви. Этот эффект может оказаться преобладающим при низких температурах.

Результирующая термо-ЭДС, состоящая из трех рассмотренных составляющих, зависит от разности температур спаев термоэлемента и от электрофизических свойств полупроводниковых материалов, образующих ветви термоэлемента. В небольшом диапазоне температур термо-ЭДС $\mathscr{E}$ можно считать (с достаточной для практических целей точностью) пропорциональной разности температур спаев термоэлемента и некоторому коэффициенту $\alpha_{1,2}$, называемому *коэффициентом термо-ЭДС:*

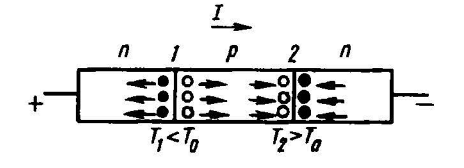
$$\mathscr{E} = \alpha_{1,2}\Delta T . \qquad (13.1)$$

Поглощение и выделение теплоты в спаях термоэлемента (эффект Пельтье)

При прохождении постоянного тока через термоэлемент в его спаях поглощается или выделяется в зависимости от направления тока некоторое количество теплоты, пропорциональное времени, току и *коэффициенту Пельтье:*

$$Q_П = \pm П_{1,2} I t \,. \qquad (13.2)$$

Если направление тока в цепи с термоэлементами такое, как показано на рис. 13.4, то на спае *1* свободные электроны и дырки, возникшие в результате тепловой генерации, движутся в разные стороны под действием суммарного электрического поля (диффузионного поля спая и внешнего поля). При тепловой генерации носителей заряда в области спая *1* на переброс электрона из валентной зоны в зону проводимости затрачивается некоторая тепловая энергия кристаллической решетки полупроводника. Поэтому при заданном направлении тока спай *1* будет охлаждаться, спай *2* при этом нагревается, так как к нему с разных сторон подходят электроны и дырки, которые отдают при рекомбинации некоторую энергию в виде теплоты.



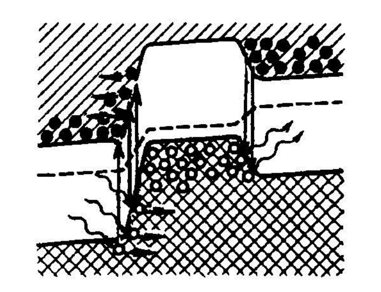


Рис. 13.4. Охлаждение спая *(1)* и нагрев спая *(2)* полупроводниковых термоэлементов при прохождении тока и энергетические диаграммы, поясняющие эти явления

В результате при прохождении тока термоэлемент работает как своеобразный тепловой насос, забирая тепловую энергию на спае *1* (теплопоглощающем спае при данном направлении тока) и выделяя ее на спае *2* (тепловыделяющем спае).

Эффект Пельтье является обратным эффекту Зеебека. Поэтому для одного и того же термоэлемента существует связь между коэффициентом Пельтье и коэффициентом термо-ЭДС:

$$П_{1,2} = \alpha_{1,2} T \,. \qquad (13.3)$$

Это соотношение может быть получено при применении к термоэлектрическим явлениям первого и второго законов термодинамики.

§ 13.2. ТЕРМОЭЛЕКТРИЧЕСКИЕ ГЕНЕРАТОРЫ

Термоэлектрнческнй генератор (ТЭГ) — это полупроводннковое термоэлектрнческое устройство, состоящее нз полупроводннкового термоэлектрнческого блока н предназначенное для непосредственного преобразовання тепловой энергнн в электрнческую энергню.

Коэффициент полезного действия термоэлемента

В соответствин с назначением одним из основиых параметров термоэлектрического геиератора является коэффициент полезного действия, т. е. отношение полезной мощности, выделяемой в иагрузке генератора, к количеству теплоты, поступающей в единицу времеии на теплопоглощающие спаи.

Рассмотрим работу отдельного термоэлемента в качестве преобразователя тепловой энергии в электрическую. Если удельиые сопротивления ветвей термоэлемента ϱ_1 и ϱ_2, *токовая высота ветви* — длина пути тока в ветви термоэлемента l, *токовое сечение ветвей* — сечение ветвей термоэлемента, определяемое по нормали к векторным лиииям тока S_1 и S_2 (см. рис. 13.1), то полное сопротивление термоэлемента

$$R = \varrho_1 l/S_1 + \varrho_2 l/S_2 \,. \qquad (13.4)$$

Переходными сопротивлениями спаев ветвей термоэлемента с металлическими контактными пластинами пренебрегаем по сравнеиию с сопротивлением ветвей термоэлемента.

Полную теплопроводность термоэлемента можно выразить через удельные теплопроводности ветвей $\varkappa_1$ и $\varkappa_2$:

$$K = \varkappa_1 S_1/l + \varkappa_2 S_2/l \,. \qquad (13.5)$$

В результате процесса теплопроводиости от теплопоглощающего спая или от источника тепловой энергии к тепловыделяющему спаю в единицу времени передается теплота

$$Q_К = К(T_2 - T_1) \,. \qquad (13.6)$$

Используя выражеиие (13.1), для тока, проходящего в цепи термоэлемеита, запишем

$$I = \frac{\alpha_{1,2}(T_2 - T_1)}{R + R_н} \,. \qquad (13.7)$$

При этом в нагрузке выделяется полезная мощность

$$P_н = I^2 R_н = \frac{\alpha_{1,2}^2 (T_2 - T_1)^2 R_н}{(R + R_н)^2} \,. \qquad (13.8)$$

Ток, проходящий в цепи термоэлемента и обусловленый эффектом Зеебека, вызывает, во-первых, выделение и поглощение

теплоты из-за эффекта Пельтье на спаях термоэлемента и, во-вторых, выделение джоулевой теплоты в ветвях термоэлемента.

Направление тока в термоэлементе соответствует поглощению в единицу времени на теплопоглощающем спае теплоты

$$Q_{П2} = П_{1,2}I = \alpha_{1,2}IT_2 . \qquad (13.9)$$

Эффект Пельтье, заключающийся в перекачивании некоторого количества теплоты от теплопоглощающего спая к тепловыделяющему, как бы увеличивает теплопроводность термоэлемента. Теплота, отдаваемая тепловыделяющему спаю в результате этого эффекта,

$$Q_{П1} = \alpha_{1,2}IT_1 . \qquad (13.10)$$

Из-за разных значений T_2 и T_1 значения $Q_{П2}$ и $Q_{П1}$ также различны. Вычитая (13.10) из (13.9) и учитывая (13.1), получим

$$Q_{П2} - Q_{П1} = \alpha_{1,2}I(T_2 - T_1) = \mathscr{E}I = I^2(R + R_н) . \qquad (13.11)$$

Таким образом, вся электрическая мощность, вырабатываемая термоэлементом, — это разница теплоты, отданной источником теплопоглощающему спаю, и теплоты, перенесенной на тепловыделяющий спай из-за эффекта Пельтье. Одна часть этой электрической мощности ($I^2R_н$) выделяется в виде полезной мощности в нагрузке, другая часть (I^2R) — в виде джоулевой теплоты в ветвях термоэлемента. В первом приближении можно считать, что джоулева теплота распределяется пополам между спаями, т. е. к теплопоглощающему спаю или источнику тепловой энергии возвращается мощность

$$\frac{1}{2}Q_{Дж} = \frac{1}{2} I^2R = \frac{\alpha_{1,2}^2(T_2 - T_1)^2 R}{2(R + R_н)^2} . \qquad (13.12)$$

Определив все количество теплоты, поступающее в единицу времени на теплопоглощающий спай термоэлемента от источника тепловой энергии, и учтя часть джоулевой теплоты, возвращенную теплопоглощающему спаю при прохождении тока по цепи термоэлемента, коэффициент полезного действия термоэлемента можно записать следующим образом:

$$\eta = \frac{P_н}{Q_К + Q_{П2} - \frac{1}{2}Q_{Дж}} . \qquad (13.13)$$

Если учесть (13.6), (13.8), (13.9) и (13.12), то

$$\eta = \frac{\dfrac{\alpha_{1,2}^2(T_2 - T_1)^2 R_н}{(R + R_н)^2}}{K(T_2 - T_1) + \dfrac{\alpha_{1,2}^2(T_2 - T_1)}{R + R_н} - \dfrac{\alpha_{1,2}^2(T_2 - T_1)^2 R}{2(R + R_н)^2}} . \qquad (13.14)$$

Обозначим отношение сопротивлений

$$R_н/R = m\,.$$

Тогда

$$R_н = mR = ml(\varrho_1/S_1 + \varrho_2/S_2) \tag{13.15}$$

и

$$R + R_н = (m+1)R = (m+1)l(\varrho_1/S_1 + \varrho_2/S_2)\,. \tag{13.16}$$

Используя (13.5), (13.15) и (13.16), преобразуем (13.14):

$$\eta = \frac{T_2 - T_1}{T_2}\,\frac{m/(m+1)}{1 + \frac{\mathrm{K}R(m+1)}{\alpha_{1,2}^2 T_2} - \frac{T_2 - T_1}{2(m+1)T_2}}. \tag{13.17}$$

Первый множитель в (13.17) представляет собой коэффициент полезного действия обратимой тепловой машины, второй множитель характеризует уменьшение КПД вследствие необратимых потерь на теплопроводность и джоулеву теплоту в термоэлементе.

Произведение $\mathrm{K}R$ в знаменателе (13.17) зависит от параметров материала ветвей термоэлемента ϱ_1 и ϱ_2, $\varkappa_1$ и $\varkappa_2$, а также от токовых сечений ветвей термоэлемента S_1 и S_2 [см. (13.4) и (13.5)]. Для обеспечения наибольшего КПД (при данных T_2, T_1, $\alpha_{1,2}$, m, ϱ и $\varkappa$ ветвей) необходимо выбрать токовые сечения ветвей так, чтобы произведение $\mathrm{K}R$ было минимальным. Для этого продифференцируем

$$\mathrm{K}R = (\varkappa_1 S_1 + \varkappa_2 S_2)\left(\frac{\varrho_1}{S_1} + \frac{\varrho_2}{S_2}\right) = \left(\frac{\varkappa_1 S_1}{S_2} + \varkappa_2\right)\left(\frac{\varrho_1 S_2}{S_1} + \varrho_2\right) \tag{13.18}$$

по (S_1/S_2) и приравняем нулю производную. Функция (13.18) имеет минимальное значение при оптимальном отношении токовых сечений ветвей

$$(S_1/S_2)_{опт} = \sqrt{\varrho_1\varkappa_2/(\varrho_2\varkappa_1)}\,. \tag{13.19}$$

При этом

$$(\mathrm{K}R)_{\min} = (\sqrt{\varkappa_1\varrho_1} + \sqrt{\varkappa_2\varrho_2})^2\,. \tag{13.20}$$

Величину $\alpha_{1,2}^2/\mathrm{K}R$, обратную величине, входящей в знаменатель выражения (13.17), с учетом оптимального отношения токовых сечений ветвей (13.19) обычно обозначают Z и называют *эффективностью* или *добротностью термоэлемента*:

$$Z = \frac{\alpha_{1,2}^2}{(\sqrt{\varkappa_1\varrho_1} + \sqrt{\varkappa_2\varrho_2})^2}\,. \tag{13.21}$$

Таким образом, КПД термоэлемента оптимальных размеров зависит от следующих факторов: 1) эффективности термоэлемен-

та Z, определяемой только электрофизическими параметрами полупроводниковых материалов ветвей термоэлемента; 2) разности температур спаев термоэлемента; 3) отношения сопротивления нагрузки к сопротивлению термоэлемента.

Для нахождения максимального КПД необходимо выбрать оптимальное отношение сопротивления нагрузки к сопротивлению термоэлемента. Дифференцируя выражение для КПД (13.17) по m и приравнивая нулю производную, получим

$$m_{опт} = (R_н/R)_{опт} = \sqrt{1 + Z\bar{T}}, \qquad (13.22)$$

где $\bar{T} = (T_2 - T_1)/2$ — средняя температура спаев.

Подставляя (13.21) и (13.22) в общую формулу (13.17), находим максимальное значение КПД термоэлемента, зависящее лишь от температуры спаев и эффективности термоэлемента:

$$\eta_{max} = \frac{T_2 - T_1}{T_2} \frac{m_{опт} - 1}{m_{опт} + T_1/T_2}. \qquad (13.23)$$

Из соотношений (13.23) и (13.22) видно, что при определенных температурах T_2 и T_1 и при $Z \to \infty$ максимальный КПД стремится к КПД идеальной тепловой машины (цикл Карно): $\eta_{max} \to (T_2 - T_1)/T_2$. Однако для приближения максимального КПД термоэлемента к КПД идеальной тепловой машины необходимо большое значение эффективности термоэлемента Z.

Часто от термоэлемента необходимо получить *максимальную мощность* даже в ущерб КПД. При таком режиме работы сопротивление нагрузки должно быть равно сопротивлению термоэлемента, т. е. $R_н/R = m = 1$. При этом

$$\eta|_{m=1} = \frac{1}{2} \frac{T_2 - T_1}{T_2 + \frac{2}{Z} - \frac{1}{4}(T_2 - T_1)} \qquad (13.24)$$

В большинстве случаев максимальный КПД (13.23) и КПД в режиме максимальной отдаваемой мощности (13.24) отличаются друг от друга на 2...3% (относительных).

Основные требования к материалам для термоэлементов

В соответствии с (13.23) и (13.24) для термоэлементов целесообразно использовать материалы, обеспечивающие большое значение эффективности Z и выдерживающие наибольшую температуру T_2, точнее — с наибольшим значением $\bar{Z}\bar{T}$ ($\bar{Z}$ — среднее значение эффективности материала в диапазоне температур от T_1 до T_2).

Как видно из (13.21), эффективность термоэлемента является функцией коэффициента термо-ЭДС $\alpha_{1,2}$, удельных сопротивлений ϱ_1 и ϱ_2 и удельных теплопроводностей $\varkappa_1$ и $\varkappa_2$ материалов ветвей термоэлемента. Все эти параметры зависят от концентрации

носителей заряда (электронов или дырок). Удельное сопротивление обратно пропорционально концентрации носителей заряда (рис. 13.5). Удельная теплопроводность слагается из теплопроводности кристаллической решетки $\varkappa_р$ и электронной теплопроводности $\varkappa_{эл}$: $\varkappa = \varkappa_р + \varkappa_{эл}$. Удельная теплопроводность решетки в первом приближении не зависит от концентрации носителей заряда, а удельная электронная теплопроводность пропорциональна их концентрации.

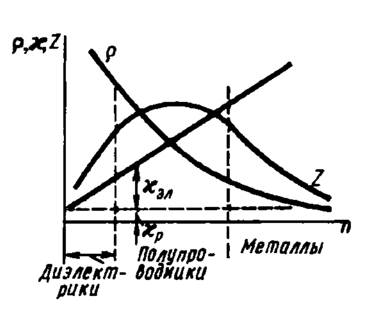


Рис. 13.5. Зависимости удельного сопротивления ρ, теплопроводности ϰ и эффективности термоэлемента Z от концентрации носителей заряда

У термоэлементов из металлов или сплавов (термопары) значение эффективности Z мало из-за низкого коэффициента термо-ЭДС и больших значений удельной теплопроводности. У термоэлементов из диэлектрических материалов эффективность Z мала из-за больших значений удельного сопротивления. У термоэлементов из материалов со средней концентрацией носителей заряда, что соответствует полупроводниковым материалам, эффективность достигает максимального значения. Поэтому полупроводниковые термоэлементы экономичнее металлических термопар. Максимум эффективности Z, как показывают расчеты, лежит в диапазоне концентраций носителей заряда (2...3) 10^{19} см$^{-3}$. Такие концентрации на три порядка меньше концентрации носителей заряда в металлах.

Конструкция, свойства и параметры полупроводниковых термобатарей

Коэффициент термо-ЭДС большинства термоэлементов ($\alpha_{1,2} = \alpha_1 - \alpha_2$) составляет 300...500 мкВ/К. Учитывая, что почти все термоэлектрические генераторы работают при разности температур спаев, не превышающей 600 К, от одного термоэлемента можно получить термо-ЭДС 0,3 В. При этом, согласно (13.24), оптимальные условия получения электрической мощности на нагрузке достигаются при сопротивлении нагрузки, сравнимом с сопротивлением термоэлемента, т. е. около нескольких миллиом. Для получения приемлемых напряжений и для согласования с реальными нагрузочными сопротивлениями необходимо последовательное соединение многих термоэлементов.

В большинстве термоэлектрических генераторов (ТЭГ) полупроводниковая термобатарея представляет собой плоскую конструкцию, состоящую из нескольких десятков термоэлементов (рис. 13.6). По форме термоэлементы могут быть разнообразными в зависимости от конструктивных особенностей источника

тепловой энергии. Наиболее распространены термоэлементы из ветвей в виде параллелепипедов с различным соотношением токового сечения и токовой высоты.

Соединение отдельных ветвей между собой металлическими контактными пластинами должно быть сделано так, чтобы контактные сопротивления были намного меньше сопротивления ветвей термоэлемента. Иначе контактные сопротивления могут уменьшить эффективность термоэлемента из-за увеличения

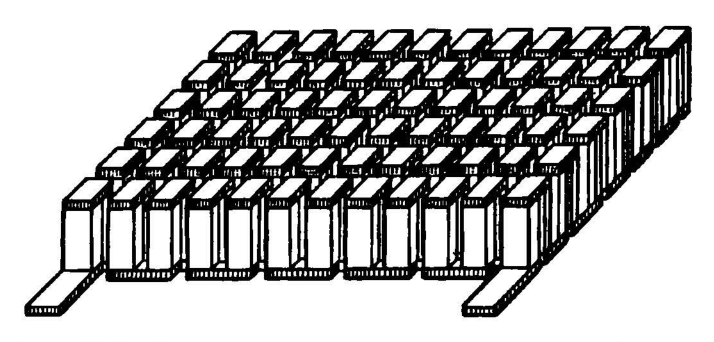


Рис. 13.6. Один из вариантов конструкции термобатареи

общего сопротивления термоэлемента при неизменном значении термо-ЭДС.

В качестве источников тепловой энергии в ТЭГ применяют атомные реакторы, источники теплоты с радиоактивными изотопами, используют химические реакции горения и тепловую энергию, излучаемую Солнцем.

В настоящее время существуют три основные области применения термоэлектрических генераторов: 1) для катодной защиты магистральных газо- и нефтепроводов от коррозии; 2) для питания электрической энергией космических аппаратов; 3) для питания электрической энергией навигационных и метеорологических установок.

§ 13.3. ХОЛОДИЛЬНИКИ И ТЕПЛОВЫЕ НАСОСЫ

Максимальное снижение температуры теплопоглощающих спаев термоэлемента

При прохождении тока через термоэлементы наряду с поглощением на спае *1* (см. рис. 13.4) в единицу времени некоторого количества теплоты

$$Q_{\Pi} = -\Pi_{1,2} I \qquad (13.25)$$

происходит еще выделение джоулевой теплоты, пропорциональ-

ной квадрату силы тока и сопротивлению спая 1. В первом приближении можно считать, что сопротивления спаев 1 и 2 равны, т. е. каждое из них составляет половину полного сопротивления термоэлемента R. Тогда джоулева теплота, выделяющаяся на первом спае,

$$Q_{Дж} = \frac{1}{2} I^2 R . \quad (13.26)$$

Алгебраическая сумма соотношений (13.25) и (13.26) для теплопоглощающего спая

$$Q = Q_{П} + Q_{Дж} = -\Pi_{1,2} I + \frac{1}{2} I^2 R . \quad (13.27)$$

Графическое пояснение уравнений (13.27) дано на рис. 13.7. При токе, равном нулю, оба эффекта отсутствуют и охлаждения не происходит, а при достаточно большом токе $Q_{Дж}$ превысит $Q_{П}$ и охлаждение первых спаев перейдет в нагрев. Поэтому существует оптимальный ток, при котором эффект охлаждения максимален (от теплопоглощающего спая отводится максимальное количество теплоты). Дифференцированием (13.27) находим, что Q достигает максимума при оптимальном токе

$$I_{опт} = \Pi_{1,2}/R . \quad (13.28)$$

При этом токе

$$Q_{I опт} = -\Pi_{1,2}^2/(2R) . \quad (13.29)$$

Рис. 13.7. Зависимость теплоты, выделяемой $Q_{Дж}$ и поглощаемой $Q_{П}$ на спае 1 термоэлемента (см. рис. 13.4), а также результирующая зависимость теплоты Q от значения тока через термоэлемент

Из (13.29) следует, что чем меньше сопротивление термоэлемента, тем больше количество теплоты $Q_{I опт}$, отводимое от теплопоглощающего спая. Однако из этого не следует, что, увеличивая токовые сечения или уменьшая токовую длину ветвей термоэлемента, можно достичь более низких температур на теплопоглощающем спае, так как одновременно в такой же мере возрастают теплопроводность и поток теплоты по ветвям термоэлемента от тепловыделяющего спая к теплопоглощающему. Температура теплопоглощающего спая будет понижаться до тех пор, пока количество теплоты $Q_К$, переносимое между спаями в результате теплопроводности по ветвям термоэлемента, не сбалансирует отводимое от спая 1 количество теплоты Q_I при токе I. (В данном случае не учитываем поток теплоты из окружающей среды на теплопоглощающий спай термоэлемента, т. е. предполагаем, что теплоизоляция теплопоглощающего спая идеальна.)

Условием теплового баланса, следовательно, будет

$$-Q_l = Q_к = К(T_2 - T_1)$$

или

$$T_2 - T_1 = -Q_l/К.$$

В оптимальном режиме работы термоэлемента, т. е. с учетом (13.29), получим

$$(T_2 - T_1)_{max} = -Q_{l\,max}/К = \Pi_{1,2}^2/(2КR). \qquad (13.30)$$

Используя (13.3) и (13.21), находим

$$(T_2 - T_1)_{max} = \frac{\alpha_{1,2}^2}{КR}\frac{1}{2}T_1^2 = \frac{1}{2}ZT_1^2. \qquad (13.31)$$

Таким образом, требования, которым должен удовлетворять полупроводник ветвей термоэлемента, работающего в холодильнике, сводятся к достижению возможно большего значения эффективности Z (как и для термоэлемента, работающего в термоэлектрическом генераторе).

Применение полупроводниковых холодильников

Полупроводниковые холодильники успешно применяют в радиоэлектронике, медицине, сельском хозяйстве, метрологии и для бытовых целей (например, автомобильные холодильники).

Однако экономичность полупроводниковых холодильников при охлаждении объемов более десятков литров пока еще ниже экономичности существующих компрессионных машин. Поэтому полупроводниковые холодильники могут конкурировать с компрессионными только при охлаждении малых объемов.

Использование принципа полупроводникового теплового насоса

В полупроводниковой термобатарее при прохождении через нее постоянного тока происходит охлаждение одних спаев и нагрев других из-за поглощения и выделения теплоты на соответствующих спаях в связи с эффектом Пельтье. Это явление, во-первых, оказывается удобным для использования в термостатах: простое изменение направления тока позволяет либо охлаждать, либо нагревать спаи термоэлементов, находящиеся внутри термостата. Во-вторых, разогрев тепловыделяющих спаев в полупроводниковой термобатарее происходит не только из-за выделения обычной джоулевой теплоты, но и в результате переноса теплоты Пельтье

с теплопоглощающих спаев на тепловыделяющие. Таким образом, на тепловыделяющих спаях может выделяться большее количество теплоты, чем было затрачено электрической энергии. Так, при разности температур 10 К на каждый ватт выделяемой на тепловыделяющих спаях теплоты необходимо затратить не более 0,2 Вт электрической энергии. Принцип теплового насоса может оказаться экономически целесообразным даже для обогрева и охлаждения жилых и производственных помещений. Но это будет возможно только при существенном снижении стоимости термоэлементов, когда выигрыш в расходе электрической энергии компенсирует стоимость полупроводникового термоэлектрического подогревателя.

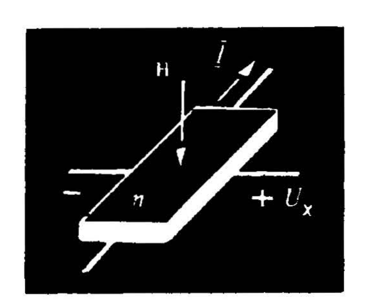
Контрольные вопросы

1. По каким причинам в полупроводниковых термоэлементах может возникать термо-ЭДС?

2. Что необходимо для получения максимального КПД термоэлектрического генератора при непосредственном преобразовании тепловой энергии в электрическую?

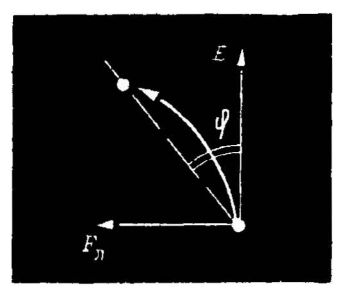
3. Каков принцип действия полупроводниковых холодильников?

4. Как можно использовать эффект Пельтье для подогрева и почему этот эффект не используют для обогрева жилых и производственных помещений?



Глава 14

Полупроводниковые гальваномагнитные приборы



§ 14.1. ПРИНЦИП ДЕЙСТВИЯ

Полупроводниковые гальваномагнитные приборы — это полупроводниковые приборы, в которых используется воздействие магнитного поля на движущиеся в электрическом поле носители зарядов.

Принцип действия различных полупроводниковых гальваномагнитных приборов основан на двух эффектах: на *эффекте Холла,* т. е. на эффекте возникновения поперечной разности потенциалов в полупроводнике, по которому проходит электрический ток, в том случае, когда есть магнитное поле, перпендикулярное направлению тока, и на *магниторезистивном эффекте,* т. е. на эффекте изменения электрического сопротивления полупроводника под действием магнитного поля *(эффект Гаусса)*. Оба эффекта вызваны тем, что на движущийся со скоростью **v** носитель заряда в магнитном поле с индукцией **B** действует сила Лоренца

$$\mathbf{F}_л = q[\mathbf{vB}]. \qquad (14.1)$$

Рассмотрим подробнее эти эффекты.

Эффект Холла

Предположим, что в пластинке полупроводника, находящейся в магнитном поле, идет ток, обусловленный движением только электронов (рис. 14.1, *а*). Пренебрежем пока статистическим разбросом электронов по скоростям. Тогда снла Лоренца будет смещать движущиеся электроны к левой грани пластинки полупроводника. Направление смещения определяется направлением силы Лоренца, т. е. векторным произведением (14.1) с учетом

знака носителей, или правилом левой руки, относящимся к техническому направлению тока. В результате смещения движущихся электронов между боковыми гранями пластинки полупроводника возникает ЭДС Холла.

В полупроводнике с электропроводностью p-типа при том же техническом направлении тока вектор скорости дырок направлен противоположно вектору скорости электронов, знак носителей

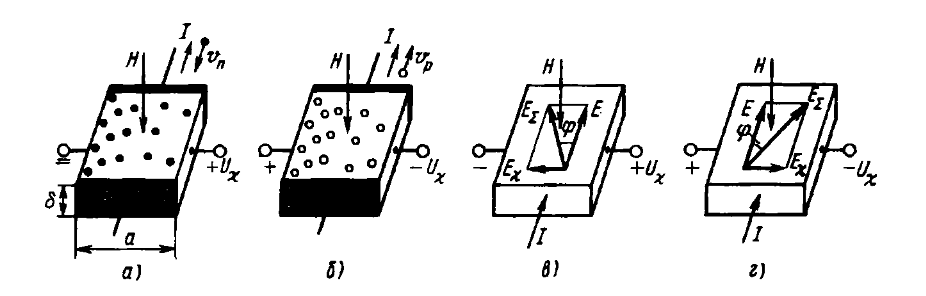


Рис. 14.1. Возникновение поперечного электрического поля при прохождении электрического тока через полупроводник, помещенный в магнитное поле:
а, в — для полупроводника с электропроводностью *n*-типа; *б, г* — для полупроводника с электропроводностью *p*-типа

заряда также другой. Поэтому сила Лоренца (14.1) действует на дырки в ту же сторону, смещая их также к левой грани пластинки полупроводника (рис. 14.1, *б*). Полярность ЭДС Холла при этом получается другой.

Накопление носителей заряда у боковой грани пластинки полупроводника прекратится, когда сила Лоренца уравновесится силой холловского электрического поля. При перпендикулярном направлении напряженности магнитного поля к поверхности пластинки полупроводника условием такого динамического равновесия будет равенство

$$qvB = qE_X. \qquad (14.2)$$

Считая холловское электрическое поле однородным и учитывая геометрические размеры пластинки полупроводника, запишем для ЭДС Холла, т. е. для поперечной разности потенциалов между боковыми гранями пластинки полупроводника с электропроводностью p-типа

$$\mathcal{E}_X = E_X a = vBa.$$

Значение скорости дырок определим из формулы для тока:

$$I = JS = qp\mu_p ES = qpva\delta \ .$$

Тогда

$$\mathscr{E}_X = \frac{1}{qp}\frac{IB}{\delta} = X\frac{IB}{\delta}, \qquad (14.3)$$

где $X = 1/(qp)$ — *коэффициент Холла* для полупроводника с электропроводностью p-типа.

В действительности носители заряда в полупроводнике распределены по скоростям. Это распределение зависит от преобладающего механизма рассеяния носителей в конкретном полупроводнике. Поэтому более точное значение коэффициента Холла отличается от имеющегося в выражении (14.3) множителем A:

$$X = A/(qp). \qquad (14.4)$$

Значение множителя A находится в диапазоне от 1 до 2 и зависит от механизма рассеяния носителей заряда. Так, для вырожденного полупроводника $A = 1$, для полупроводника с преобладающим рассеянием носителей на тепловых колебаниях кристаллической решетки $A = 1{,}18$, для полупроводника с преобладающим рассеянием на ионизированных примесях $A = 1{,}93$.

Для полупроводника с электропроводностью n-типа полярность ЭДС Холла противоположна. Поэтому коэффициент Холла для такого полупроводника имеет другой знак:

$$X = -A/(qn).$$

В полупроводниках с приблизительно равными концентрациями электронов и дырок (например, в собственных полупроводниках) расчет коэффициента Холла получается более сложным:

$$X = -\frac{A}{q}\frac{n\mu_n^2 - p\mu_p^2}{(n\mu_n + p\mu_p)^2}.$$

После возникновения холловской напряженности электрического поля и установления динамического равновесия между силой Лоренца и силой холловского электрического поля все носители заряда, имеющие скорость v, будут двигаться по прямолинейным траекториям в соответствии с направлением внешнего электрического поля $\mathbf{E}$ (рис. 14.1, *в, г*). При этом направление вектора суммарного электрического поля ($\mathbf{E}_\Sigma = \mathbf{E} + \mathbf{E}_X$) отличается от технического направления вектора тока на некоторый угол φ (рис. 14.1, *в. г*), который называют *углом Холла*. Угол Холла определяют по формуле

$$\operatorname{tg}\varphi = E_X/E.$$

Холловская напряженность электрического поля в полупроводнике с электропроводностью p-типа с учетом (14.3)

$$E_X = \frac{1}{qp}\frac{IB}{\delta a} = \frac{1}{qp}JB.$$

Напряженность в пластинке полупроводника от внешнего источника питання

$$E = J/\gamma = J/(qp\mu_p).$$

Поэтому

$$\operatorname{tg}\varphi = E_X/E = \mu_p B. \qquad (14.5)$$

Очевндно, что для полупроводниковой пластинки с электропроводностью n-тнпа получнтся аналогнчное соотношенне между углом Холла, подвижностью электронов и зиачеинем магннтной иидукцин. При малых магнитиых полях и, следовательно, при малых углах Холла

$$\operatorname{tg}\varphi \approx \varphi,$$

нли

$$\varphi \approx \mu B. \qquad (14.6)$$

Магниторезистивиый эффект

Под действием снлы Лоренца траектория движеиия носителя заряда искривляется, что равносильно уменьшению длины свободиого пробега в направлеиии внешнего поля между токовыми контактамн или увеличенню удельного сопротивления полупроводника в магнитном поле.

Эффект Холла мы рассматривали на примере пластиики полупроводника, помещенной в магнитное поле. В этом случае при установлении дннамнческого равновесия возникшая холловская иапряжениость электрического поля компенсирует действие силы Лоренца и, следовательно, не происходит искрнвлеиия траекторнй носителей заряда, движущихся со скоростью v. Казалось бы, в таком случае сопротивление полупроводинка ие должно изменяться под действнем магиитного поля. Однако эти рассуждеиия справедливы только для носнтелей, движущихся со скоростью v, соответствующей средней скорости. В действительности же, как было отмечено, носителн в полупроводннке распределены по скоростям. Поэтому носители со скоростью, превышающей среднюю скорость, смещаются к одной граии пластиикн полупроводника, так как иа них действует бóльшая сила Лореица (14.1). Носители, обладающие скоростью, меиьшей средией скорости, смещаются к другой грани пластннки полупроводника, так как на ннх действует бóльшая сила холловской напряжениости электрнческого поля. Такнм образом, в рассмотренном прнмере удельное сопротивленне полупроводника изменяется в магнитиом поле из-за нскривления траекторий носителей заряда, движущнхся со скоростью, отличиой от средней скоростн.

Холловская напряженность электрнческого поля частично или полиостью компеисирует действие силы Лореица в зависнмостн от скорости иосителей заряда. Поэтому иаибольший магиторе-

зистивный эффект можно получить в полупроводнике таких формы и конструкции, при которых возникновение холловской напряженности электрического поля будет затруднено или вообще невозможно. Теоретически такую конструкцию можно представить, если предположить неограниченность полупроводниковой пластинки в направлении, параллельном токовым контактам. Практическая возможность реализации таких условий рассмотрена в § 14.3.

В неограниченном полупроводнике не происходит накопления носителей заряда на боковых гранях пластинки полупроводника, не образуется ЭДС Холла, а траектория движения носителей заряда отклоняется от направления внешнего электрического поля в направлении силы Лоренца (рис. 14.2). В этом случае вектор плотности тока совпадает по направлению со скоростью движения дырок и поэтому оказывается сдвинутым от направления внешнего электрического поля на тот же угол φ — угол Холла.

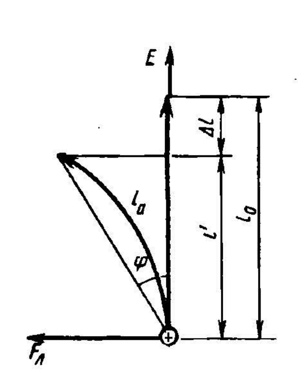


Рис. 14.2. Изменение длины свободного пробега носителя заряда (дырки) вдоль вектора электрического поля в полупроводнике, находящемся в магнитном поле

Рассмотрим теперь магниторезистивный эффект с количественной точки зрения при оптимальных условиях, т. е. для неограниченного полупроводника.

Отклонение траектории движения носителей заряда от направления внешнего электрического поля в неограниченном полупроводнике равносильно уменьшению длины свободного пробега носителей заряда в направлении электрического поля на

$$\Delta l = l_0 - l' = l_0 - l_0 \cos \varphi,$$

где l_0 — длина свободного пробега носителей заряда при отсутствии магнитного поля; l' — проекция пути, пройденного носителем заряда между двумя последовательными столкновениями (длина свободного пробега) при наличии магнитного поля, на направление внешнего электрического поля (рис. 14.2).

При малых магнитных полях и, следовательно, при малых значениях угла Холла φ можно воспользоваться разложением в ряд:

$$\cos \varphi \approx 1 - \varphi^2/2! + \ldots .$$

Тогда

$$\Delta l \approx l_0 - l_0 + l_0 \varphi^2/2.$$

Используя (15.6), получим

$$\Delta l \approx (l_0/2) \mu^2 B^2.$$

Так как за время свободного пробега носитель заряда проходит при наличии магнитного поля меньший путь вдоль электрического поля E, то это эквивалентно уменьшению дрейфовой скорости или подвижности, а следовательно, и удельной проводимости полупроводника. Относительное изменение удельного сопротивления при этом

$$(\varrho-\varrho_0)/\varrho_0=(\gamma_0-\gamma)/\gamma_0=\Delta l/l_0=\mu^2B^2/2. \qquad (14.7)$$

Для ограниченного по размерам кристалла полупроводника обычно пользуются соотношением

$$\Delta\varrho/\varrho_0=C\mu^2B^2,$$

где C — коэффициент, зависящий от геометрических размеров пластинки полупроводника.

§ 14.2. ПРЕОБРАЗОВАТЕЛИ ХОЛЛА

Преобразователь Холла — это гальваномагнитный полупроводниковый прибор, основанный на использовании эффекта Холла.

Технология изготовления и конструкция

Выражение (14.3) справедливо для пластинки полупроводника, длина которой значительно больше ширины. ЭДС преобразователей Холла конечной длины зависит от отношения l/a:

$$\mathcal{E}_{\text{X}}=\text{X}\frac{IB}{\delta}f\left(\frac{l}{a}\right). \qquad (14.8)$$

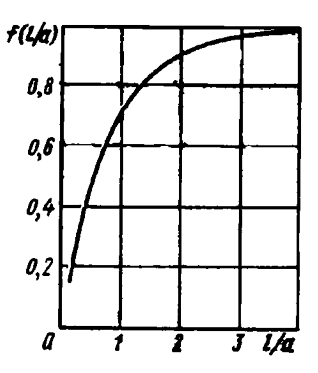


Рис. 14.3. Зависимость поправочной функции к ЭДС Холла от отношения длины и ширины преобразователя Холла

Результаты расчета поправочной функции $f(l/a)$ приведены на рис. 14.3. Как видно из рисунка, поправка к ЭДС Холла, связанная с конечной длиной преобразователя, составляет всего несколько процентов при $l/a=2...3$. Поэтому нет смысла сильно увеличивать длину преобразователя Холла. Такое соотношение длины и ширины имеют практически все преобразователи Холла.

Преобразователи Холла могут быть изготовлены в виде пластинок прямоугольной формы, вырезанных из монокристалла полупроводника (например, германия) или поликристаллического полупроводника, или в виде тонких (1...10 мкм) пленок, нанесенных на диэлектрическую подложку методом испарения в вакууме. В качестве материала тонких пленок используют такие соединения, как селенид и теллурид

ртути HgSe, HgTe и их твердые растворы. Эти соединения обладают большим значением подвижности электронов [около 10^4 см^2/(В·с)] даже при малых толщинах пленок. Для изготовления преобразователей Холла применяют также сурьмянистый индий InSb — соединение типа $A^{III}B^{V}$, имеющее рекордно высокую подвижность электронов, и другие материалы.

К пластинке полупроводника припаивают четыре электрода: два входных (токовых) и два выходных (холловских). Выходные контакты преобразователей Холла обычно не удается располо-

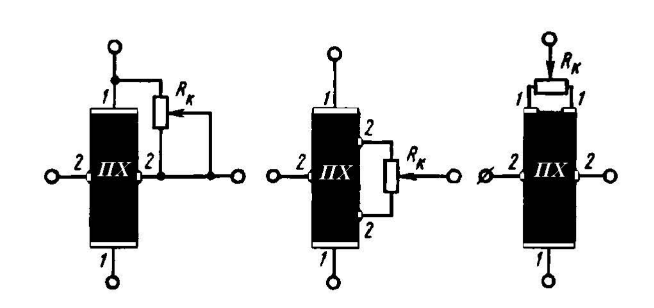


Рис. 14.4. Схемы для компенсации неэквипотенциальности выходных электродов:

ПХ — преобразователь Холла; *1* — входные электроды; *2* — выходные электроды; $R_К$ — компенсационный резистор

жить строго на эквипотенциальной линии (при $B=0$) из-за возможной несимметрии выходных контактов или из-за неоднородности полупроводникового материала. К искажению эквипотенциалей и в результате к неэквипотенциальности выходных контактов приводят также механические напряжения, которые могут возникнуть в различных частях пластинки полупроводника при изготовлении преобразователя Холла. В этом случае при прохождении тока и при отсутствии магнитного поля на выходных электродах будет некоторая разность потенциалов, возникновение которой не связано с эффектом Холла. Эта разность потенциалов устраняется с помощью компенсирующих схем (рис. 14.4).

Искажения показаний преобразователя Холла при работе в переменном магнитном поле могут быть также вызваны ЭДС индукции в контуре выходных электродов. Значение этого паразитного сигнала можно уменьшить, если скрутить вместе провода, идущие от выходных электродов, т. е. уменьшить площадь контура, пронизываемую магнитным потоком. ЭДС, наводимую в контуре выходных электродов, можно скомпенсировать с помощью дополнительной петли, помещенной в то же магнитное поле (рис. 14.5).

Основные параметры и свойства

Входным сопротивлением преобразователя Холла называют сопротивление между входными электродами. При отсутствии магнитного поля его можно определить по удельному сопротивлеиию исходного материала и по геометрическим размерам:

$$R_{вх}=\varrho l/(a\delta). \qquad (14.9)$$

Сопротивление полупроводников, в том числе входиое сопротивление преобразователя Холла, измеияется в магиитиом поле, возрастая при увеличеиии иапряжеииости магиитиого поля.

Первой причиной изменения входного сопротивления преобразователя Холла, помещениого в магиитиое поле, является измеиеиие подвижности иосителей заряда (см. § 14.1). Вторая причииа — изменеиие концентрации носителей заряда из-за влияиия поверхиостной рекомбинации носителей, которые отклоияются силой Лоренца к боковой поверхиости преобразователя.

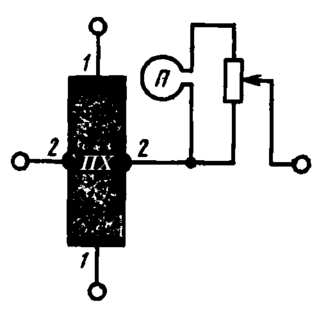


Рис. 14.5. Схема компенсации паразитной ЭДС, индуцируемой в цепи выходных электродов:

ПХ — преобразователь Холла; 1 — входные электроды; 2 — выходные электроды; П — петля для компенсации ЭДС индукции в цепи выходных электродов

Выходным сопротивлением преобразователя Холла называют сопротивлеиие между выходными электродами. Если выходиые контакты точечиые, то выходиое сопротивление определяется удвоеиным сопротивлением растекания точечиого контакта (см. § 3.21):

$$R_{вых}\approx\varrho/(\pi a_к), \qquad (14.10)$$

где $a_к$ — радиус точечиого выходиого коитакта.

При выходиых коитактах, выполиенных в виде полуцилиидра, вплавленных в боковые грани полупроводниковой пластиики, выходиое сопротивление преобразователя Холла

$$R_{вых}=\frac{2\varrho}{\pi\delta}\ln\frac{a}{2a_к},$$

где $a_к$ — радиус полуцилиидрического выходиого контакта.

При иаличии тока в цепи холловских электродов выходное сопротивление преобразователя Холла, так же как и входиое сопротивление, возрастает с увеличеиием напряжеиности магиитиого поля.

Коэффициент использования или КПД преобразователя Холла определяется отиошением отдаваемой и подводимой мощиостей:

$$\eta=P_н/P_{вх}.$$

Мощность, подводимая к преобразователю Холла, с учетом (14.9)

$$P_{вх}=I^2\frac{\varrho l}{a\delta}. \qquad (14.11)$$

Мощность, выделяемая на сопротивлении нагрузки,

$$P_н=I_X^2R_н=\frac{U_X^2R_н}{(R_{вых}+R_н)^2}.$$

При согласовании выходного сопротивления преобразователя Холла и сопротивления нагрузки, т. е. при равенстве этих сопротивлений ($R_{вых}=R_н$), преобразователь отдает в нагрузку максимальную мощность при фиксированной подводимой мощности. Тогда для преобразователя с точечными выходными контактами

$$P_н=U_X^2\pi a_к/(4\varrho).$$

Коэффициент использования такого преобразователя Холла с учетом соотношений (14.3) и (14.4)

$$\eta=\frac{\pi a a_к}{4\varrho^2\delta l}X^2B^2=\left(\frac{A}{2}\right)^2\mu^2B^2\frac{\pi a a_к}{\delta l}.$$

При условии согласования сопротивления нагрузки и выходного сопротивления для преобразователя Холла с полуцилиндрическими выходными контактами и с учетом выражения (14.11) запишем

$$P_н=\frac{\pi\delta U_X^2}{8\varrho\ln\frac{a}{2a_к}},$$

коэффициент использования

$$\eta=\frac{\pi a X^2B^2}{8\varrho^2 l\ln\frac{a}{2a_к}}=\left(\frac{A}{2}\right)^2\mu^2B^2\frac{\pi a}{2l\ln\frac{a}{2a_к}}.$$

Таким образом, при различных выходных контактах коэффициент использования преобразователей Холла пропорционален квадрату подвижности носителей заряда.

Максимально допустимый ток. Для увеличения ЭДС Холла и выходной мощности необходимо увеличивать входную мощность. При подводимой к преобразователю Холла мощности, равной допустимой мощности рассеяния, температура преобразователя повышается до максимально допустимой. Поэтому

$$I_{max}^2\frac{\varrho l}{a\delta}=\beta S\Delta T, \qquad (14.12)$$

где S — площадь поверхности преобразователя Холла; β — коэффициент теплоотдачи; ΔT — разность температур между максимально допустимой и температурой окружающей среды.

Если пренебречь площадью боковых граней и считать, что $S=2al$, то на основании уравнения теплового баланса (14.12) максимально допустимый ток через преобразователь Холла

$$I_{max}=a\sqrt{2\beta\Delta T\delta/\varrho}. \qquad (14.13)$$

Согласно выражению (14.13), увеличить максимально допустимый ток, а следовательно, и входную мощность, не изменяя температуры преобразователя, можно, увеличивая интенсивность теплообмена между преобразователем и окружающей средой.

Следует отметить, что максимальная мощность рассеяния преобразователя определяется теплоотдачей с его поверхности только у очень тонких пленочных преобразователей, а также у кристаллических преобразователей, имеющих поверхность, превышающую 70...80 мм^2. У кристаллических преобразователей малых размеров даже тонкие электроды отводят теплоты больше, чем рассеивается с поверхности самого преобразователя.

Максимальная ЭДС Холла при заданной индукции магнитного поля возникает в преобразователе при прохождении через него максимально допустимого тока. Как следует из (14.3) и (14.13), максимальная ЭДС Холла

$$\mathcal{E}_{X max}=XBa\sqrt{\frac{2\beta\Delta T}{\varrho\delta}}.$$

Вольтовая чувствительность представляет собой отношение максимальной ЭДС Холла к индукции магнитного поля:

$$\gamma=\mathcal{E}_{X max}/B=aX\sqrt{2\beta\Delta T/(\varrho\delta)}.$$

При оценке качества преобразователя Холла вольтовая чувствительность является более удобным и объективным параметром, чем максимальная ЭДС Холла, которая зависит от магнитной индукции.

Вольтовую чувствительность преобразователя Холла можно значительно увеличить при использовании импульсного режима питания. Максимально допустимый ток в импульсном режиме можно увеличить в десятки и сотни раз, соответственно увеличится и максимальная ЭДС Холла при заданной индукции магнитного поля.

Коэффициент передачи равен отношению напряженности ЭДС Холла к напряженности электрического поля между входными электродами при определенной магнитной индукции. С учетом соотношений (14.3), (14.4) и (14.9) коэффициент передачи преобразователя Холла

$$K=E_X/E=|\mathrm{tg}\,\varphi|=A\mu B.$$

При рассмотрении основных параметров преобразователей Холла можно подумать, что качество преобразователя определяется только подвижностью носителей заряда в полупроводниковой пластинке. Однако это не совсем так. Действительно, при

высокой подвижности ЭДС Холла составляет существенную часть входного напряжения, т. е. коэффициент передачи велик. Но при этом само входное напряжение должно быть достаточно большим, чтобы полезный выходной сигнал имел необходимое значение. Например, для преобразователя Холла из полупроводника с электропроводностью p-типа ЭДС Холла при допустимой мощности рассеяния, т. е. с учетом выражения (14.11),

$$\mathscr{E}_X = \frac{A}{qp}\frac{IB}{a} = AB\sqrt{\frac{P_{\text{вх.доп}}a}{q\delta l}}\sqrt{\frac{\mu}{p}}. \quad (14.14)$$

Поэтому ЭДС на выходе преобразователя Холла или его вольтовая чувствительность в соответствии с (14.14) зависит не только от подвижности носителей заряда, но и от их концентрации. В антимониде индия с рекордно большой подвижностью носителей заряда концентрация носителей не может быть малой из-за малой ширины запрещенной зоны. В результате преобразователи Холла из германия обладают лучшими параметрами.

Частотные свойства. Время релаксации процессов, определяющих эффект Холла, порядка $10^{-12}...10^{-13}$ с, что и определяет теоретический частотный предел использования преобразователей Холла. Однако при их практическом использовании на очень высоких частотах встречаются определенные трудности.

В соответствии с тем, что на преобразователь Холла воздействуют две управляющие величины: магнитная индукция и входной ток, — при рассмотрении частотных свойств необходимо различать два случая.

В первом случае преобразователь Холла находится в постоянном магнитном поле, а входная цепь питается высокочастотным управляющим током. В поведении преобразователя на постоянном и переменном токе до частот в несколько десятков мегагерц не наблюдается какой-либо разницы в пределах точности измерений. Чем больше частота входного тока, тем сильнее сказывается нелинейность контактов полупроводника с металлическими электродами.

Во втором случае преобразователь Холла питается постоянным управляющим током и находится в высокочастотном магнитном поле. При таком режиме работы в пластинке полупроводника индуцируются вихревые токи, которые вызывают ее нагрев. Вследствие этого максимально допустимый ток необходимо уменьшать при использовании преобразователя Холла в высокочастотном магнитном поле.

Применение преобразователей Холла

Применение преобразователей Холла основано на том, что их выходной сигнал (ЭДС Холла) пропорционален произведению тока на магнитную индукцию. В связи с этим преобразователи

Холла можно примеиять для измереиия магнитиых полей, токов, а также в качестве множительных устройств, фазочувствительных детекторов, анализаторов спектра и т. п.

§ 14.3. МАГНИТОРЕЗИСТОРЫ

Магниторезистор — это полупроводниковый резистор, в котором используется зависимость электрического сопротивления от магнитного поля.

Принцип действия магниторезисторов осиоваи на магиторезистивиом эффекте (см. § 14.1). Поскольку холловская напряженность электрического поля, возникающая в полупроводнике с током при наличии магнитного поля, снижает магниторезистивный эффект, то конструкция магниторезистора должна быть такой, чтобы уменьшить или полностью устранить ЭДС Холла. Наилучшей формой магиторезистора является *диск Корбино* (рис. 14.6, *а*).

При отсутствии магнитного поля ток в таком магниторезнсторе проходит в радиальном направлении от центра диска ко второму электроду, расположенному по периметру диска, или наоборот. Под действием магнитного поля носители заряда отклоняются в направлении, перпендикулярном радиусу. Так как не существует граней, на которых может происходить накопление зарядов, то ЭДС Холла в таком магниторезисторе не возникает.

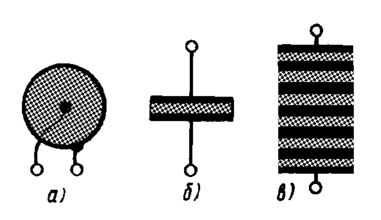


Рис. 14.6. Различные коиструкции магииторезисторов:

а — магниторезистор в виде диска Корбино; *б* — магниторезистор, изготовленный из пластинки полупроводиика с малой длииой; *в* — магииторезистор из пластиики или плеики полупроводника с поперечными металлическими полосками на поверхности

Другой конструкцией магниторезистора является пластннка полупроводника, ширина которой много больше ее длины (рис. 14.6, *б*). Однако существенным недостатком магниторезистора такой конструкции является его малое сопротивление, для увеличения которого применяют последовательное соединение нескольких магниторезисторов или нанесение на поверхность пластины полупроводника металлических полос (рис. 14.6, *в*). Каждая часть пластины полупроводника между двумя металлическими полосами представляет собой отдельный магниторезистор. Можно также считать, что металлические полосы выполняют роль шунтов, уменьшающих ЭДС Холла, возникающую на боковых гранях пластины полупроводника.

Основным полупроводниковым материалом для магниторезисторов является антимонид индия InSb и арсенид индия InAs — материалы с большой подвижностью носителей заряда.

§ 14.4. МАГНИТОДИОДЫ И МАГНИТОТРАНЗИСТОРЫ

Магнитодиоды

Магнитодиод — это полупроводниковый диод, в котором используется изменение вольт-амперной характеристики под действием магнитного поля.

Для диодов с тонкой базой ($W<L$) и с несимметричным *p-n*-переходом (например, p^+-*n*) формулу тока насыщения запишем в соответствии с (3.31) и с учетом (1.29):

$$I_{нас}=S\frac{p_{n0}\mu_p kT}{W_n}. \qquad (14.15)$$

Следовательно, сопротивление диода с тонкой базой под действием магнитного поля изменяется только в результате изменения подвижности неосновных носителей заряда в базе диода. Из сравнения выражений (14.15) и (14.7) видно, что изменение тока, проходящего через диод с тонкой базой, или изменение его сопротивления в магнитном поле значительно меньше, чем в магниторезисторе. Таким образом, диоды с тонкой базой нецелесообразно использовать в качестве магниточувствительных полупроводниковых приборов — магнитодиодов.

В диодах с толстой базой ($W>L$) прямое напряжение, приложенное к диоду, распределяется между *p-n*-переходом и сопротивлением базы диода:

$$U_{пр}=U_{p\text{-}n}+I_{пр}R_{б}.$$

Сопротивление базы диода увеличивается в поперечном магнитном поле в результате уменьшения подвижности основных и неосновных носителей заряда, как и в обычном магниторезисторе. Увеличение сопротивления базы диода с толстой базой может быть связано также с уменьшением времени жизни неосновных носителей, если из-за искривления траектории движения неосновные носители будут достигать поверхности базовой области, где велика скорость их рекомбинации.

В результате увеличения сопротивления базы прямое напряжение, приложенное к диоду с толстой базой, перераспределяется: уменьшается доля напряжения, приходящаяся на *p-n*-переход. Этот процесс приводит к резкому уменьшению тока, проходящего через диод, так как этот ток связан экспоненциальной зависимостью с напряжением на *p-n*-переходе. Кроме того, ток, проходящий через *p-n*-переход, уменьшается из-за уменьшения тока насыщения, как и в диоде с тонкой базой.

Таким образом, диод с толстой базой может быть использован в качестве магнитодиода при соответствующем выборе геометрических размеров базы диода и электрофизических параметров исходного материала. Обычно магнитодиоды изготовляют с толщиной базы, соответствующей нескольким диффузионным длинам неосновных носителей, т. е. толщиной в несколько мил-

лиметров. Полупроводииковый материал базы, так же как и для магниторезисторов, должен обладать большой подвижностью носителей заряда [см. соотношение (14.7)]. Этим требованиям удовлетворяют германий и кремний.

Прямые ветви ВАХ германиевого магнитодиода в магнитных полях с различной магнитной индукцией показаны на рис. 14.7. Для оценки чувствительности магнитодиода к магнитному полю, по аналогин с преобразователями Холла, используют вольтовую чувствительность.

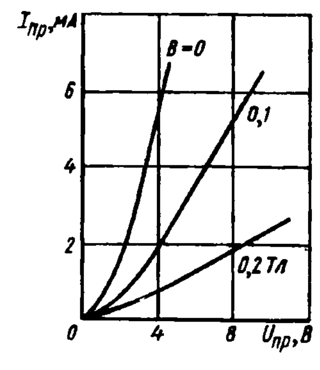


Рис. 14.7. Прямые ветви ВАХ германиевого магнитодиода, находящегося в магнитных полях с различной магнитиой индукцией

$$\gamma = \Delta U/(BI),$$

где ΔU — изменение напряжения на магнитодиоде при внесеиии его в магнитное поле.

Вольтовая чувствительность магнитодиодов может быть значительно выше вольтовой чувствительности преобразователей Холла из того же материала.

Биполяриые магиитотраизисторы

Магнитотраизистор — это траизистор, в котором используется зависимость его характеристик и параметров от магнитного поля.

Обычно биполярные транзнсторы малочувствительны к магнитному полю, так как поперечное магнитное поле приводит только к искривленню траекторий движения неосновных носителей заряда, идущих через базу от эмиттера к коллектору, что эквиваленто уменьшению эффективной подвижности неосновных носителей в базе транзистора. В связи с малой толщиной базы в обычных биполярных транзисторах практически все инжектированные эмиттером носители достигают коллектора, несмотря на искривление траекторий их двнжения магнитным полем. Другой физической причиной изменения параметров биполярных транзисторов в магнитном поле является изменение сопротивления базы транзистора.

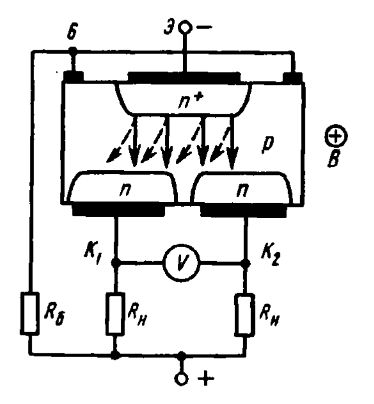


Рис. 14.8. Структура биполярного магнитотраизистора с двумя коллекторами и схема его включения

Для увеличення чувствительно-

сти к магнитному полю биполярные магнитотранзисторы делают с двумя коллекторными переходами (рис. 14.8). Без магнитного поля половина инжектированных носителей заряда попадает на один коллектор, половина — на другой. Магнитное поле отклоняет носители от одного коллектора к другому. По изменению токов первого и второго коллекторов можно оценить или измерить магнитную индукцию поперечного магнитного поля, используя для этого, например, своеобразную мостовую схему (рис. 14.8).

Биполярные магнитотранзисторы в диапазоне слабых магнитных полей могут иметь магнитную чувствительность, на несколько порядков большую магнитной чувствительности преобразователей Холла.

Контрольные вопросы

1. В чем заключается эффект Холла?
2. Что такое угол Холла и от чего он зависит?
3. Что такое магниторезистивный эффект?
4. Почему вольтовая чувствительность преобразователей Холла, изготовленных из материала с большой подвижностью носителей заряда, оказывается иногда меньше, чем вольтовая чувствительность таких же преобразователей из материала с меньшей подвижностью носителей?
5. Какую конструкцию должны иметь магниторезисторы?
6. Какие диоды можно использовать в качестве магнитодиодов?
7. Что такое биполярный магнитотранзистор?

ЗАКЛЮЧЕНИЕ

В заключение отметим некоторые перспективы развития полупроводниковых приборов, а также проблемы, которые предстоит решить для успешного развития полупроводниковой электроники.

Одной из важнейших задач полупроводниковой электроники является увеличение рабочих частот, увеличение быстродействия полупроводниковых приборов, в том числе и интегральных микросхем. Прогресс в этом направлении достигнут значительный: максимальная частота генерации биполярных транзисторов за тридцать с лишним лет, прошедших после появления первых плоскостных транзисторов, возросла на несколько порядков и достигла 10 ГГц. Значение этого параметра биполярных СВЧ-транзисторов уже близко к теоретическому пределу. Перечислим фундаментальные физические ограничения, определяющие теоретический предел быстродействия различных полупроводниковых приборов.

Первым из них является конечность времени релаксации заряда, т. е. времени установления электрической нейтральности различных частей структуры полупроводникового прибора. Время релаксации $\tau=\varepsilon\varepsilon_0\varrho$ должно быть значительно меньше полупериода переменного сигнала. Это необходимо для того, чтобы за время изменения входного напряжения на биполярном транзисторе успела измениться высота потенциального барьера эмиттерного перехода, в полевом транзисторе — успело произойти изменение толщины канала, в варикапе — изменилась бы толщина *p-n*-перехода, в генераторе Ганна — сформировался бы домен и т. д.

С другой стороны, максимальная концентрация примесей в базе диодов, биполярных транзисторов или в подложке полевых транзисторов ограничена сверху обычно значениями 10^{16}... 10^{17} см$^{-3}$. Время релаксации заряда для таких концентраций примесей в кремнии получается 10^{-12}...$3\cdot10^{-13}$ с.

Вторым фундаментальным ограничением быстродействия полупроводниковых приборов является конечность скоростей движения носителей заряда и соответственно определенное время, необходимое для отбора энергии от постоянного электрического

поля носителем заряда. Минимальное время изменения энергии электрона на величину kT

$$t_{kT} = kT/(qE_{max}v_{max}),$$

где E_{max} — максимально допустимая напряженность электрического поля, выше которой наступает пробой; v_{max} — максимальная скорость дрейфа электронов.

Для кремния при $T=300$ K, $E_{max}=3\cdot10^5$ В/см, $v_{max}=10^7$ см/с получается $t_{kT}=0,9\cdot10^{-14}$ с. В большинстве полупроводниковых приборов процесс усиления — отбор энергии от постоянного электрического поля и передача части энергии переменному электрическому полю — происходит в p-n-переходе, где напряженность электрического поля изменяется с координатой. Поэтому часть пути носитель заряда испытывает воздействие напряженности электрического поля, значительно меньшее E_{max}. То же самое можно сказать о скорости носителей заряда. Таким образом, реальное время, необходимое для приобретения носителем заряда добавочной энергии в несколько kT, должно быть более 10^{-12} с.

Кроме этих двух фундаментальных физических причин ограничения быстродействия в различных полупроводниковых приборах необходимо учитывать постоянные времени перезаряда барьерных емкостей p-n-переходов (в биполярных транзисторах, тиристорах, диодах, полевых транзисторах с управляющим переходом), постоянные времени перезаряда распределенных емкостей в МДП-структурах МДП-транзисторов и приборов с зарядовой связью.

Таким образом, для продвижения вверх по частотному диапазону наряду с совершенствованием различных полупроводниковых СВЧ-приборов (генераторов Ганна, лавинно-пролетных диодов, биполярных и полевых транзисторов) необходимо изыскание новых принципов усиления и генерации электрических колебаний. Увеличение быстродействия выпрямительных диодов (особенно мощных, что очень важно) может быть достигнуто путем использования гетеропереходов и выпрямляющих электрических переходов между металлом и полупроводником, т. е. структур без инжекции неосновных носителей заряда в базу диода. При этом можно исключить относительно медленный процесс накопления неосновных носителей и соответственно процесс рассасывания этих носителей.

Другой проблемой является повышение допустимой мощности рассеяния полупроводниковых приборов, что трудно осуществить не в ущерб быстродействию этих приборов. Проблема отвода теплоты характерна не только для мощных выпрямительных диодов, транзисторов, генераторов Ганна, полупроводниковых лазеров и других дискретных полупроводниковых приборов, но и для интегральных микросхем.

Одним из решений этой проблемы может быть использование полупроводниковых материалов с большой шириной запрещенной зоны (большей, чем у кремния). Собственная проводимость у

широкозонных полупроводников становится существенной при более высоких температурах. Поэтому приборы на их основе могут работать при бóльших удельных мощностях рассеяния. Одним из таких материалов является арсенид галлия. Одновременно с отмеченным преимуществом арсенида галлия следует также учесть еще бóльшую подвижность носителей заряда в нем, что должно обеспечить увеличение быстродействия полупроводниковых приборов на основе арсенида галлия как в дискретном, так и в интегральном исполнении.

Однако для разработки интегральных микросхем на основе арсенида галлия необходимо преодолеть еще ряд технологических трудностей: в частности, научиться наращивать на монокристаллы арсенида галлия диэлектрические слои, не создавая при этом большой плотности поверхностных состояний. Именно поэтому до сих пор не удалось изготовить качественные МДП-транзисторы на арсениде галлия.

Важнейшими задачами, всегда существующими в полупроводниковой электронике, являются повышение надежности и уменьшение стоимости дискретных и интегральных полупроводниковых приборов. Для решения этих задач также необходимо всестороннее изучение физических процессов, происходящих в полупроводниковых приборах, совершенствование существующих технологий их производства, разработка новых технологических приемов, а также изыскание новых принципов действия и эффектов, с помощью которых можно выполнять необходимые функциональные преобразования.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

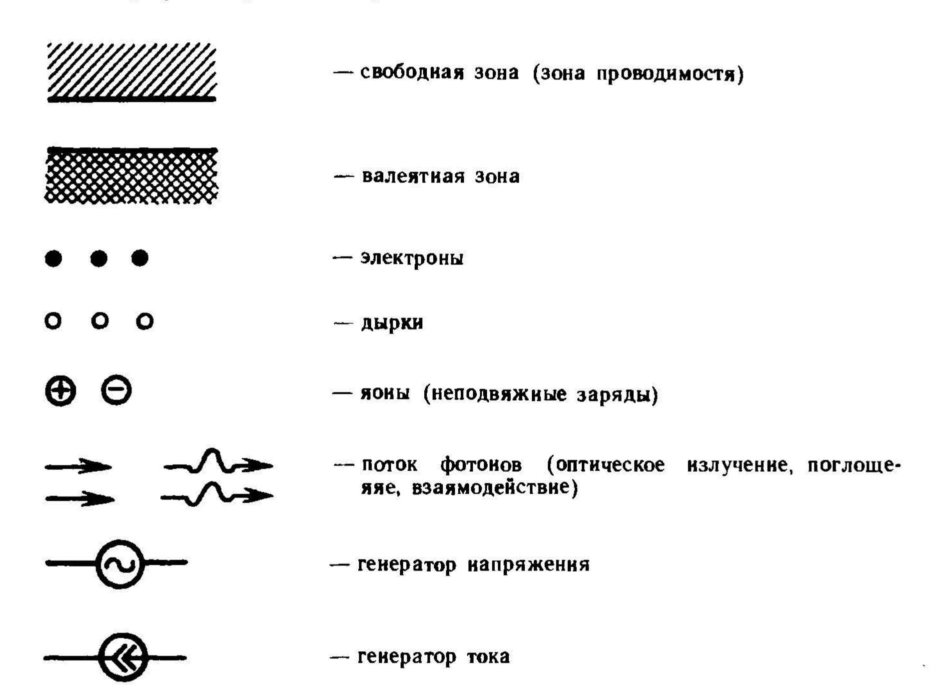
1. Степаненко И. П. Основы теории транзисторов и транзисторных схем. — М.: Энергия, 1977.
2. Батушев В. А. Электронные приборы. — М.: Высшая школа, 1980.
3. Викулин И. М., Стафеев В. И. Физика полупроводниковых приборов. — М.: Советское радио, 1980.
4. Зи С. Физика полупроводниковых приборов: Пер. с англ. — М.: Мир, 1984.
5. Морозова И. Г. Физика электронных приборов. — М.: Атомиздат, 1980.
6. Степаненко И. П. Основы микроэлектроники. — М.: Советское радио, 1980.
7. Агаханян Т. М. Интегральные микросхемы. — М.: Энергоатомиздат, 1983.
8. Терехов В.А. Задачник по электронным приборам. — М.: Энергоатомиздат, 1983.
9. ГОСТ 22622—77 «Материалы полупроводниковые. Термины и определения основных электрофизических параметров».
10. ГОСТ 15133—77 (СТ СЭВ 2767—80) «Приборы полупроводниковые. Термины и определения».
11. ГОСТ 2730—73 (СТ СЭВ 661—77) «Единая система конструкторской документации. Обозначения условные графические в схемах. Приборы полупроводниковые».
12. ГОСТ 18472—82 (СТ СЭВ 1818—79) «Приборы полупроводниковые. Основные размеры».
13. ГОСТ 19613—80 «Столбы и блоки выпрямительные полупроводниковые. Основные размеры».
14. ГОСТ 25529—82 (СТ СЭВ 1125—78, СТ СЭВ 2768—80) «Диоды полупроводниковые. Термины, определения и буквенные обозначения параметров».
15. ГОСТ 20003—74 (СТ СЭВ 2770—80) «Транзисторы биполярные. Термины, определения и буквенные обозначения параметров».
16. ГОСТ 20332—84 (СТ СЭВ 1125—78) «Тиристоры. Термины, определения и буквенные обозначения параметров».
17. ГОСТ 19095—73 (СТ СЭВ 2771—80) «Транзисторы полевые. Электрические параметры. Термины, определения и буквенные обозначения».
18. ГОСТ 17021—75 (СТ СЭВ 1623—79) «Микросхемы интегральные. Термины и определения».
19. ГОСТ 22274—80 (СТ СЭВ 3787—82) «Излучатели полупроводниковые. Термины, определения и буквенные обозначения параметров».
20. ГОСТ 24354—80 «Приборы полупроводниковые визуального представления информации. Основные размеры».
21. ГОСТ 21934—83 «Приемники излучения полупроводниковые фотоэлектрические и устройства фотоприемные. Термины и определения».
22. ГОСТ 23562—79 (СТ СЭВ 3787—82) «Оптопары. Термины, определения и буквенные обозначения параметров».
23. ГОСТ 24403—80 (СТ СЭВ 3787—82) «Микросхемы интегральные оптоэлектронные. Термины, определения и буквенные обозначения параметров».
24. ГОСТ 18577—80 «Устройства термоэлектрические полупроводниковые. Термины и определения».

ОБОЗНАЧЕНИЯ ОСНОВНЫХ ВЕЛИЧИН, ПРИНЯТЫЕ В КНИГЕ

B — магнитная индукция, коэффициент температурной чувствительности, яркость
C — емкость
D — коэффициент диффузии
E — напряженность электрического поля, освещенность
$\mathcal{E}$ — ЭДС
F — коэффициент шума
f — частота
g — плотность объемного заряда
h — постоянная Планка ($\hbar = h/2\pi$)
I — постоянное, эффективное значение тока, интенсивность света
i — мгновенное значение тока
J — плотность тока
j — $\sqrt{-1}$
k — постоянная Больцмана, волновое число
L — диффузионная длина носителей заряда, индуктивность
l — длина, расстояние, длина свободного пробега носителя заряда
M — коэффициент лавинного размножения
m — масса, тензочувствительность
m^* — эффективная масса носителя заряда
N — концентрация примесей, эффективная плотность энергетических уровней
n — концентрация электронов
n_0 — равновесная концентрация электронов
P — постоянное, эффективное значение мощности, импульс
p — концентрация дырок
p_0 — равновесная концентрация дырок
Q — заряд, добротность, количество теплоты
q — элементарный заряд
Φ — световой поток
R — статическое сопротивление, радиус
r — дифференциальное сопротивление, радиус
S — площадь, крутизна характеристики
s — скорость поверхностной рекомбинации
T — абсолютная температура, период колебаний
ТКI — температурный коэффициент тока
ТКR — температурный коэффициент сопротивления
ТКU — температурный коэффициент напряжения
t — температура, время
U — постоянное, эффективное значение напряжения
u — мгновенное значение напряжения
V — объем
v — скорость
W — толщина базы
$Э$ — энергия

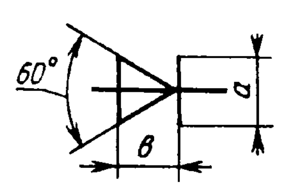
α — коэффициент передачи тока эмиттера одномерной теоретической модели транзистора, коэффициент ударной ионизации, коэффициент термо-ЭДС, показатель поглощения
α^* — эффективность коллектора
α_n — коэффициент переноса
β — коэффициент передачи тока базы одномерной теоретической модели транзистора, коэффициент нелинейности
γ — удельная проводимость, эффективность эмиттера
δ — толщина *p-n*-перехода
ε — относительная диэлектрическая проницаемость
ε_0 — электрическая постоянная
η — коэффициент полезного действия
ϰ — удельная теплопроводность
λ — длина волны
μ — подвижность
μ_n — коэффициент Пуассона
ν — частота в оптическом диапазоне
Π — коэффициент Пельтье
ρ — удельное сопротивление
σ — проводимость
τ — время жизни носителей заряда, постоянная времени
φ — потенциал, угол, фазовый сдвиг
X — коэффициент Холла
ω — угловая частота

На рисунках приняты следующие обозначения:

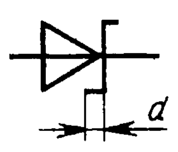


УСЛОВНЫЕ ГРАФИЧЕСКИЕ ОБОЗНАЧЕНИЯ ОСНОВНЫХ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ В СХЕМАХ

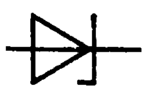
Диод



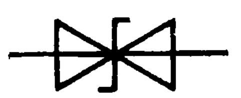
Диод Шотки



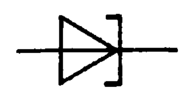
Стабилятрон (диод лавинный выпрямительный)



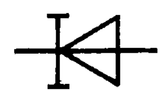
Стабилятрон двуханодный



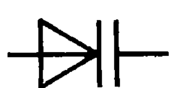
Туннельный диод



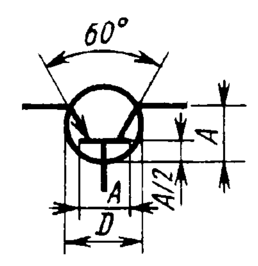
Обращенный диод



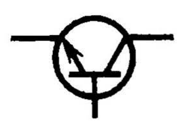
Варикап



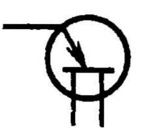
Биполярный транзистор типа *p-n-p*



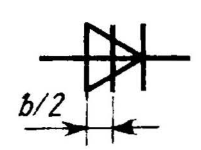
Биполярный транзистор типа *n-p-n* .



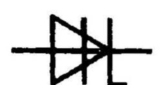
Однопереходный транзистор с *n*-базой



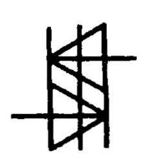
Диодный тиристор



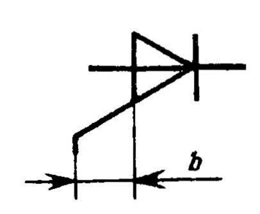
Диодный тиристор, проводящий в обратном направлении



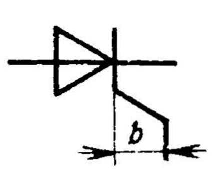
Диодный симметричный тиристор



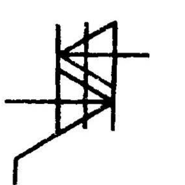
Триодный тиристор с управлением по аноду



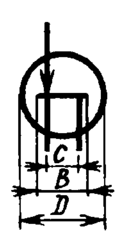
Триодный тиристор с управлением по катоду



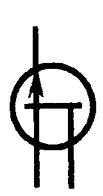
Триодный симметричный тиристор



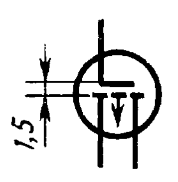
Полевой транзистор с управляющим переходом и n-каналом

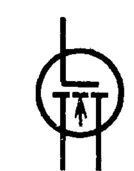


Полевой транзистор с управляющим переходом и p-каналом

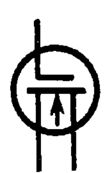


Полевой транзистор с изолированным затвором обогащенного типа (с индуцированным p-каналом и n-каналом)

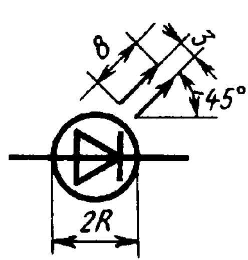




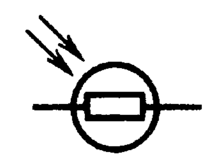
Полевой транзистор с изолированным затвором обедненного типа (со встроенным p-каналом и n-каналом)



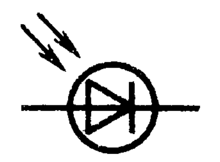
Излучающий диод



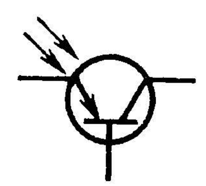
Фоторезистор



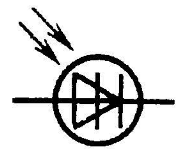
Фотодиод



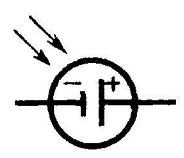
Фототранзистор типа *p-n-p*



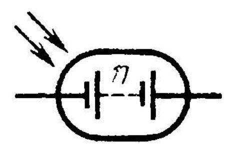
Фототиристор



Фотоэлемент



Солнечная фотобатарея



Размеры (в миллиметрах) условных графических обозначений полупроводниковых приборов в соответствии с ГОСТ 2700—73 (СТ СЭВ 661—77) могут иметь следующие значения:

A	*a*	*B*	*b*	*c*	*D*	*d*	*R*
9	5	8	4	6	12	1,5	5
11	6	9	5	7	14	2	6

УНИВЕРСАЛЬНЫЕ ФИЗИЧЕСКИЕ ПОСТОЯННЫЕ

Постоянная Планка

$$h \approx 6{,}63 \cdot 10^{-34}\ \text{Дж} \cdot \text{с} = 4{,}14 \cdot 10^{-15}\ \text{эВ} \cdot \text{с}.$$

Постоянная Больцмана

$$k \approx 1{,}38 \cdot 10^{-23}\ \text{Дж/К} = 0{,}86 \cdot 10^{-4}\ \text{эВ/К}.$$

Масса покоя электрона

$$m \approx 9{,}1 \cdot 10^{-31}\ \text{кг} = 9{,}1 \cdot 10^{-28}\ \text{г}.$$

Заряд электрона (элементарный заряд)

$$q \approx 1{,}6 \cdot 10^{-19}\ \text{Кл}.$$

Электрическая постоянная

$$\varepsilon_0 \approx 8{,}86 \cdot 10^{-12}\ \text{Ф/м} = 8{,}86 \cdot 10^{-14}\ \text{Ф/см}.$$

Магнитная постоянная

$$\mu_0 \approx 1{,}257 \cdot 10^{-6}\ \text{Гн/м} = 1{,}257 \cdot 10^{-8}\ \text{Гн/см}.$$

Скорость света в вакууме

$$c \approx 3 \cdot 10^{8}\ \text{м/с} = 3 \cdot 10^{10}\ \text{см/с}.$$

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

В предметный указатель включены основные термины и понятия, встречающиеся в учебнике. Рядом с термином указаны только те страницы, на которых можно найти определение или расшифровку данного термина, а также сведения, поясняющие его смысл.

Термины, состоящие из прилагательного и существительного, помещены в предметном указателе в большинстве случаев с инверсией, т. е. за основное слово принято существительное. Однако некоторые термины, используемые в технической литературе обычно без инверсии (например, «электронно-дырочный переход»), приведены в алфавитном указателе дважды.

ОГЛАВЛЕНИЕ

Оглавление